PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : A61K 35/76, C12N 7/00, 7/01, 15/09, 15/85, 15/86 — A1

(11) International Publication Number: WO 98/05344

(43) International Publication Date: 12 February 1998 (12.02.98)

(21) International Application Number: PCT/US97/12928

(22) International Filing Date: 3 July 1997 (03.07.97)

(30) Priority Data:
08/693,865 5 August 1996 (05.08.96) US
08/814,859 11 March 1997 (11.03.97) US

(71) Applicant (*for all designated States except US*): BRIGHAM AND WOMEN'S HOSPITAL, INC. [US/US]; 75 Francis Street, Boston, MA 02115 (US).

(72) Inventors; and
(75) Inventors/Applicants (*for US only*): SARKAR, Saumyendra, N. [IN/US]; 102 Queensbury Street, Boston, MA 02115 (US). KUPPER, Thomas, S. [US/US]; 8 Surrey Lane, Weston, MA 02193 (US). DUBIN, Daniel, B. [US/US]; Apartment D-618, 99 Pond Avenue, Brookline, MA 02146 (US).

(74) Agent: PLUMER, Elizabeth, R.; Wolf, Greenfield & Sacks, P.C., 600 Atlantic Avenue, Boston, MA 02210 (US).

(81) Designated States: AU, CA, JP, US, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(54) Title: BACTERIOPHAGE-MEDIATED GENE THERAPY

(57) Abstract

An improved method for delivering an exogenous gene, e.g., a therapeutic polynucleotide, to a mammalian cell is provided. The method involves using a bacteriophage as a vector to deliver the exogenous gene to a pre-selected target cell. The bacteriophage contains exogenous genetic material that can be transcribed and, optionally, translated in a mammalian cell and includes on its surface a ligand that binds to a receptor on the target cell. The bacteriophage is incapable of injecting the exogenous genetic material into the mammalian cell. The bacteriophages are useful for gene therapy applications and for producing exogenous gene products *in vitro*.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

BACTERIOPHAGE-MEDIATED GENE THERAPY

Government Support

This work was funded in part by the National Institutes of Health under the Grant Numbers AI 25062. The Government may retain certain rights in this invention.

Related Applications

This application is a continuation-in-part of U.S. Serial No. 08/814,859, filed March 11, 1989, pending, which is a continuation-in-part of U.S. Serial No. 08/693,865, filed August 5, 1996, pending, the entire contents of which are incorporated herein by reference.

Field Of The Invention

This invention relates to the use of a bacteriophage to express an exogenous gene in a mammalian cell. In particular, the invention relates to the use of bacteriophage vectors for human gene therapy and compositions related thereto.

Background Of The Invention

The application of gene therapy for the treatment of human disease has increased steadily since the first human gene therapy trial was conducted in 1989. To date, well over 100 (Science 269: 1050-5 (1995)) gene therapy protocols and clinical trials have been approved by the Recombinant DNA Advisory Committee for the treatment of inherited and acquired diseases. Despite the reported advances in gene therapy technology and the increasing approvals of gene therapy protocols by the National Institutes of Health, the delivery and long-term expression of exogenous genes in specific tissues for the treatment of genetic disease remains a formidable challenge.

Vectors which are being studied in gene therapy trials include retroviruses, adenoviruses, adeno-associated virus, plasmids and liposomes. Each has significant limitations. Although the adenovirus reportedly is a relatively efficient gene delivery vehicle for both dividing and non-dividing cells, its drawbacks include (1) limitation of inserted expression cassette to seven kilobases, (2) induction of inflammation in mammalian hosts, (3) toxicity to target cells when administered in high titers, (4) inability to target specific cell types, and (5) transient expression in infected cells. Moreover, both injection of fully packaged adenovirus and subsequent expression of adenoviral antigenic proteins in infected cells provoke a host immune response that may significantly limit the bioavailability of repeated administration of the vector. Indeed, an inflammatory response to adenovirus during a cystic fibrosis gene therapy trial has been reported.

As an alternative to adenovirus, adenoassociated virus has been proposed as an

alternative to adenovirus, and does have certain advantages, including a preferential site of chromosomal integration. Although the adeno-associated virus has not been associated with inflammation, this virus' genome is only 4.7 kilobases in size and can accommodate inserts of up to 5 kb in size. Thus, its primary limitation is in the size of the "expression cassette" that it can deliver to the target cell. An "expression cassette" is a term of art that refers to an inserted exogenous DNA that optionally contains transcriptional regulatory sequences, translational regulatory sequences, coding sequences, or intervening sequences necessary for efficient expression of the included coding sequence. The cost of manufacturing adeno-associated virus is high compared to that of adenovirus. Furthermore, adeno-associated virus is difficult to grow at high titer and cannot be targeted to defined cellular targets.

The majority of gene therapy trials employ retroviral vectors for delivering an exogenous gene into mammalian cells and are susceptible to complement-mediated destruction. However, currently utilized retroviral vectors can transduce only dividing mammalian cells. In addition, the retroviruses can accommodate inserts up to only seven kilobases, thereby limiting the size of exogenous genetic material that can be delivered using this type of vector. In addition, the cost of manufacturing retroviral vectors has been estimated to be approximately $100,000 per lot, reportedly due to the high costs associated with producing high titers of recombinant retroviral vectors in animal cells (D. Holzman, "Gene Therapy Depends on Finding the Right Vector", JNCI, Issue 6, vol. 87 (1995) NEWS pg 406). As late as 1995, only about half of the 76 retroviral clinical trials (Science 269: 1050-5 (1995)) approved for gene therapy had been initiated, in part, due to the high expense and difficulty in undertaking the manufacture of the retroviral vectors. The concern that retroviruses may replicate *in vivo* has inhibited clinical acceptance of retrovirus-mediated gene therapy and at least one incidence of replication of a retroviral vector occurring in clinical materials has been reported (D. Holzman, *ibid.*). There is also concern that random integration of retroviral vectors could disrupt or otherwise adversely affect host cell gene expression.

Non-viral vectors for gene therapy that are being studied in clinical trials include liposomes and naked DNA delivery systems. Low efficiency of transfection hampers both of these methods, neither of which permit targetable delivery to specific cell types. In addition, liposomes can be toxic to cells.

Despite the wide range of vectors currently available for human gene therapy, it is generally agreed that a clinically efficient and cost-effective vector for delivery of an exogenous

gene to specific mammalian cells or tissues has not been identified. The inadequacies of the above-identified vectors include: (1) induction of an immune response in the host; (2) possible deleterious recombination events; (3) limitations in the size of the expression cassette that can be inserted into the vector; (4) lack of specificity of the vector for delivering the expression cassette to a particular target cell; (5) inability to target a vector to both dividing and non-dividing cells; (6) high vector production cost, (7) low efficiency of transfection (non-viral vectors), (8) unacceptable toxicity, (9) low efficiency of expression in vivo, and (10) transience of expression of the exogenous genetic material. Accordingly it is incumbent upon scientists interested in realizing the unfulfilled promise of gene therapy to develop vectors that can overcome these shortcomings.

Summary Of The Invention

Applicants have discovered that certain well-characterized bacteriophages can be modified to contain exogenous genetic material that can be transcribed and, optionally, translated in a mammalian cell. Moreover, modification of these bacteriophages to include specific ligands and gene transcription regulatory elements permits control over the cell(s) to which the expression cassette is delivered and in which the delivered genetic material will be expressed. Furthermore, these bacteriophages can be modified to include nuclear localization and endosomal lysis signals to enhance the efficiency of transfection. Despite genetic modification, these bacteriophages preferably maintain the ability to propagate to high titer in a prokaryotic host, thus greatly facilitating production.

Thus, the invention provides compositions and methods for using these bacteriophages to introduce an exogenous polynucleotide (e.g., a therapeutic polynucleotide) into a pre-selected target cell via receptor-mediated delivery. According to one aspect of the invention, the bacteriophages are used to deliver a therapeutic polynucleotide into a mammalian cell for human gene therapy. In yet another aspect of the invention, the bacteriophages are used to deliver an exogenous polynucleotide into a mammalian cell for the production, *in vitro* or *in vivo,* of a polynucleotide transcription or translation product. According to yet other aspects of the invention, compositions containing the bacteriophages of the invention and methods for preparing and using the same to produce exogenous polynucleotide transcription and translation products *in vitro* also are provided. In certain preferred embodiments of the foregoing aspects of the invention, the bacteriophage contains exogenous genetic material that can be both transcribed and translated in a mammalian cell. In other preferred embodiments, the bacteriophage contains

exogenous genetic material that is transcribed, but not translated, in a mammalian cell (e.g., ribozyme and antisense constructs). Exemplary bacteriophages which can be modified to satisfy this criteria are provided in Table 1. (Tables 1-8 are presented at the end of the detailed description of the invention, immediately preceding the specific Examples section.) Advantageously, these bacteriophages can be propagated in prokaryotic cell(s), thereby permitting the large scale production of the bacteriophages of the invention at a relatively low cost in comparison with other gene therapy vectors. Furthermore, since the bacteriophage structural genes are controlled by prokaryotic promoters, no transcription of native bacteriophage genes will occur in the mammalian host cell. Therefore, cells transfected with a modified bacteriophage will not express foreign proteins on their cell surface and, thus, will not induce cell-mediated immunity.

The preferred bacteriophages of the invention include the lambda and p1 phages. These preferred phages, after targeting modifications, have in common the ability to deliver to a mammalian cell(s) a therapeutic polynucleotide that is between about 1-100 kilobases. In particular, the preferred phages have in common the ability to deliver to a mammalian cell(s) a therapeutic polynucleotide of large size: between about 9-23 kilobases for lambda Dash II and between about 75-100 kilobases for p1. The preferred modified lambda phages of the invention can accommodate between about 9-50 kilobases. Thus, the invention is particularly useful for delivering to a target cell, an exogenous polynucleotide expression cassette (e.g., a therapeutic polynucleotide) containing a single gene or multi-gene complexes that are greater than 9 kilobases in length, something that cannot be accomplished by currently available viral vectors.

According to one aspect of the invention, an improved method for gene therapy is provided. The improvement lies in using a bacteriophage to deliver a therapeutic polynucleotide to a target cell in a mammalian recipient. In particular, the invention provides, for the first time, a viral vector that delivers to a target cell, an exogenous polynucleotide containing up to 9 to 23 kilobases for lambda Dash II, up to 50 kilobases for lambda-derived cassettes containing only COS sequences and no packaging protein coding sequences (minimal lambda cassette), and up to 75 to 100 kilobases for p1. The bacteriophages of the invention contain on their surfaces ligands to achieve delivery of the bacteriophages into a pre-selected target cell. The pre-selected target cell contains on its surface a receptor that selectively binds to the ligand, forming a ligand-receptor complex that is internalized by the cell. Although Applicants do not wish to limit the scope of the invention to a particular mechanism, it is believed that formation of the ligand-

receptor complex induces internalization of the complex via a receptor-mediated delivery mechanism such as that utilized by native eukaryotic viruses (e.g., adenovirus) and as that described in U.S. Patent No. 5,108,921, issued to Low, et al. and U.S. Patent No. 5,166,320, issued to Wu, et al., internalization of the bacteriophage into the target cell (i.e., transfection) can be performed *in vivo* or *ex vivo*. According to another aspect of the invention the bacteriophage vector further contains a nuclear localization signal (e.g., retrovirus) (Naldini et al. Science 1996; 272:263) as well as an endosomal lysis signal (e.g., adenovirus)(Wagner et al. Proc Natl Acad Sci 1992;89:6009) which, alone or in combination, enhance the expression of the delivered polynucleotide. Alternatively, chloroquine or other inhibitors of lysosomal/endosomal enzymatic catabolism can be co-administered with the vector. The use of nuclear localization signal(s), endosomal lysis signal(s), and/or endosomal enzyme inhibitors optimize expression of the delivered polynucleotide by minimizing the likelihood of DNA degradation that may occur as the vector traffics to the nucleus. In the particularly preferred embodiments, the delivered polynucleotide is designed and constructed in accordance with standard practice to integrate into the target cell chromosome. Alternatively, the vector may be maintained episomally in the cell. Accordingly, the novel bacteriophages disclosed herein are useful for *in vivo* and *ex vivo* gene therapy, as well as for producing in culture or in animals, gene products of the therapeutic polynucleotides (e.g., transcription products such as antisense RNA or catalytic RNA (e.g., ribozyme), and translation products such as peptides and proteins).

The improved method for gene therapy is useful for introducing a therapeutic polynucleotide (e.g., a polynucleotide for treating or diagnosing a medical condition) into a target cell of a mammalian recipient. The recipient is diagnosed as having a medical condition that is treatable by administration to the recipient of the therapeutic polynucleotide or a product thereof (e.g., a transcription product such as an antisense RNA or a translation product such a peptide or protein). The improved gene therapy method involves: (1) contacting the bacteriophage with the target cell under conditions (a) to permit selective binding of a ligand on the surface of the bacteriophage to a receptor on the surface of the target cell and (b) to allow the bacteriophage to enter the target cell; and (2) allowing the target cell to live under conditions such that the therapeutic polynucleotide is transcribed therein. Exemplary medical conditions and their respective therapeutic polynucleotides (or therapeutic polynucleotide products) that are useful for treating these medical conditions are provided in Table 2 (see, also, BioWorld Financial Watch, Monday Sept. 19, 1994, pp4-10, American Health Consultants, Inc.; Gene

Therapy A Primer for Physicians, Culver KW, Ed. 1996, Mary Ann Liebert, Inc. New York). Therapeutic polynucleotides that are particularly suited for delivery using the bacteriophages of the invention are provided in Table 3. Exemplary target cells to which the bacteriophages can deliver these and other therapeutic polynucleotides are provided in Table 4. The preferred target cells express on their surfaces receptors that mediate cellular uptake of the bacteriophage vector via receptor-mediated endocytosis. More preferably, the target cells are non-phagocytic. Screening methods for identifying target cells that internalize the bacteriophage vectors by way of receptor-mediated endocytosis are disclosed in the Examples. In general, these methods assay uptake by the target cell of the bacteriophage vector in the presence or absence of one or more known inhibitors of receptor-mediated endocytosis. The invention is directed to bacteriophage vectors that enter the cell via receptor-mediated endocytosis and not by the natural phage prokaryotic transduction pathway. Accordingly, in the preferred vectors, the bacteriophage tail proteins that are required for natural phage transduction are either, absent, non-functional in a prokaryotic host, or not capable of mediating injection of genetic material into a eukaryotic host cell. The screening methods disclosed in the Examples permit the selection of bacteriophage vectors and target cells which satisfy the above-noted criteria.

Advantageously, the preferred bacteriophage vectors are abortive to lytic growth in the natural bacterial flora of the mammalian host. Such modified bacteriophage vectors are also abortive to lytic growth in Su^0 (natural E. Coli host without the amber suppressor gene mutation supE or supF, both of which code for tRNAs) bacterial host strains in vitro. In the preferred embodiments, the bacteriophage vector genome is modified so as to contain an amber mutation inserted in-frame into a bacteriophage tail protein gene, e.g. J, M, or H gene, such that in an Su^0 bacterial host the tail protein gene is truncated and non-functional and, hence, abortive to lytic growth. Advantageously, the bacteriophage vectors with tail protein amber mutations propagate only in E. Coli strains containing either the amber suppressor gene mutation supE or supF.

Alternatively, the bacteriophage vectors have temperature-sensitive tail protein mutations abortive to lytic growth at mammalian host physiologic temperature. At temperatures other than (greater or less than) the mammalian host physiologic temperatures, the tail proteins are expressed and function normally and can mediate prokaryotic cell infections in vitro. Accordingly, such modified bacteriophage vectors are packaged in vitro in cell free systems using purified protein packaging extracts and engineered cosmid vectors. In the preferred embodiments, the bacteriophage vectors have temperature-sensitive tail protein mutations such

that at temperatures other than (greater or less than) the mammalian host physiologic temperature, the tail proteins are expressed and function normally and can mediate prokaryotic cell infections; however, at mammalian physiologic temperature, the temperature sensitive mutation inactivates the wild type tail function.

In the preferred embodiments, the bacteriophage genome includes a promoter to control transcription and/or translation of the therapeutic polynucleotide in the target cell. Exemplary promoters are provided in Table 5. More preferably, the promoters are cell or tissue specific (i.e., they are functional only in particular types of cells or tissues), thereby providing an additional means for controlling expression (i.e., transcription and translation) of the therapeutic polynucleotide in the target cell. Optionally, the bacteriophage genome includes additional regulatory sequences, e.g., enhancers, for further controlling expression of the therapeutic polynucleotide in the target cell. Exemplary enhancers also are provided in Table 5. In the preferred embodiments, the enhancers are target cell specific. Optionally, the bacteriophage genome can include eukaryotic origins of replication (e.g., from mammalian chromosomes), telomeres and centromeres to permit autonomous replication of the bacteriophage genome within the target cell and segregation of the replicated genome into the target cell progeny.

The bacteriophage of the invention includes on its surface a ligand which selectively binds to a receptor on the target cell surface to form a ligand-receptor complex. The complex is internalized by the target cell, presumably by receptor-mediated endocytosis. Applicants do not intend to limit the invention to a particular internalization mechanism, other than limiting the invention to exclude the natural phage transduction pathway that, as described in the literature, involves tail protein mediated injection of the bacteriophage genome into the target cell. The ligand can be attached to the surface of the bacteriophage using, for example, chemical modification methods, (e.g., galactosylation), genetic engineering methods (e.g., inserting a sequence encoding the ligand into the bacteriophage genome, in frame, such that the ligand is expressed on the surface of the bacteriophage), specific adsorption (e.g., coating an antibody onto the surface of a bacteriophage) or a combination of genetic engineering and affinity binding methods (e.g., expressing avidin on the bacteriophage surface to form an "avidin-labeled bacteriophage" and binding a biotinylated ligand thereto). Such avidin-labeled bacteriophages also are useful as intermediates in attaching virtually any ligand to the surface of a bacteriophage, provided that following biotinylation, the ligand retains its functional activity (i.e., the ability to selectively bind to its receptor to form a ligand-receptor complex). Potential ligands include

peptide or nucleotide polymers, macromolecular aggregates, such as lipoproteins or any chemical structure, either naturally occurring, synthesized, or generated by combinatorial chemistry techniques. Exemplary receptors and preferred ligands for targeting specific cell types are provided in Table 6. The preferred ligands include: low density lipoprotein (apoprotein B100), very low density lipoprotein (apoprotein E or a single chain variable immunoglobulin gene fragment that has high affinity for the VLDL receptor), HDL (apoA1), galactose, c kit ligand, transferrin, insulin, heregulin, and RGD or RGD-containing polypeptides (cyclic RGD).

The invention also embraces ligand-labeled bacteriophages in which the ligand is an antibody (or fragment thereof) that selectively binds to an antibody receptor (e.g., an Fc receptor) on the surface of a target cell. The antibody can be attached to the surface of the bacteriophage by, for example, selectively binding an antibody to a bacteriophage surface antigen. Thus, the invention is useful for targeting the delivery of an exogenous polynucleotide to virtually any Fc receptor bearing target cell *in vivo* or *ex vivo* via Fc receptor-mediated internalization. Finally, the avidin labeled phage can be bound to a biotinylated antibody. This antibody can bind to its cognate antigen on the cell surface, an event that is followed by patching, capping, and internalization. Optionally, the bacteriophage is modified, after *in vitro* propagation and packaging, to prevent or reduce the likelihood that the bacteriophage will enter bacteria that may be endogenous to the mammalian recipient or that may be a contaminant of target cells in culture. Such modifications can take the form of attaching the ligand to those portions of the bacteriophage (e.g., the "D" protein of the phage head) in a manner to inhibit the ability of the bacteriophage to attach to and/or penetrate its bacterial host. Alternatively, the ligand can be attached to the surface of a bacteriophage at a location that is not involved in phage attachment or penetration of the bacterial host. Other procedures, such as genetic engineering/chemical modification, can be used to interfere with the function of one or more bacteriophage encoded translation products that are essential for phage attachment and/or penetration into the bacterial host. (See, e.g., C. Merril, et al., "Long-circulating bacteriophage as antibacterial agents", PNAS USA 93:3188-3192 (1996), which describes the importance played by the amino acid glutamic acid at position 158 of the capsid E protein in the insertion, *in vivo*, of bacteriophage lambda into *E. coli*).

According to yet another aspect of the invention, an avidin-labeled bacteriophage is provided. The bacteriophage has a genome in which the recombinant exogenous genetic material can be transcribed and translated in a mammalian cell. Since the native bacteriophage coding

sequences are regulated by prokaryotic promoters, in the ensuing discussion, the recombinant bacteriophage genome will refer only to the inserted genetic material whose expression is controlled by eukaryotic regulatory elements. Preferably, for bacteriophage vectors that are intended to produce a polypeptide, the recombinant bacteriophage genome can be both transcribed and translated in the mammalian cell. As mentioned above, the avidin-labeled bacteriophages are useful as intermediates in generating the ligand-labeled bacteriophages of the invention. Alternative high affinity binding pairs can be substituted for the avidin/biotin binding pair in accordance with the methods of the invention. For example, streptavidin can be substituted for avidin to form a streptavidin-labeled bacteriophage that can be allowed to bind to a biotinylated ligand to form a streptavidin-biotin complex. Alternatively, modified avidin or streptavidin may be employed.

According to yet another aspect of the invention, a method for introducing an exogenous polynucleotide into a target cell, preferably, a human cell, is provided. The method involves: (1) contacting the bacteriophage with the target cell under conditions: (a) to permit selective binding of a ligand on the surface of the bacteriophage to a receptor on the surface of the target cell and (b) to allow the bacteriophage to enter the target cell; and (2) allowing the target cell to live under conditions such that the exogenous polynucleotide is transcribed and, optionally, translated therein. As used herein, the phrase "exogenous polynucleotide" refers to a nucleic acid that is not normally present in the naturally-occurring (i.e., non-recombinant) bacteriophage. Thus, exogenous polynucleotides embrace the above-described therapeutic polynucleotides, as well as polynucleotides that are not intended for therapeutic applications (e.g., polynucleotides that encode a mammalian protein or protein complex for production of the protein or protein complex in cell culture, transcription regulatory elements, telomeres, centromeres, splice junctions, autonomous replicating sequences, recombination specific sequences). Exemplary bacteriophages and ligands that are useful in accordance with this method are described above in reference to the improved method for gene therapy. According to a particularly preferred embodiment, the method further involves the step of isolating an exogenous polynucleotide product (e.g., a transcription or translation product) from the target cell.

According to yet another aspect of the invention, a bacteriophage that is useful for practicing the above-described methods for delivering an exogenous polynucleotide (e.g., a therapeutic polynucleotide) to a target cell is provided. The bacteriophage, which can be propagated in prokaryotes, contains a recombinant genome that can be transcribed and,

optionally, translated in a mammalian cell. More preferably, the recombinant bacteriophage genome can be both transcribed and translated in the mammalian cell. The surface of the bacteriophage is modified to contain thereon a ligand that selectively binds to a receptor on the mammalian target cell.

The bacteriophages of the invention are useful in the preparation of a medicament for treating a medical condition that is treatable by administration to the mammalian recipient of the therapeutic polynucleotide or a product thereof. The bacteriophages can be placed in a pharmaceutically acceptable carrier to form a pharmaceutical composition which can be administered to the recipient in accordance with standard clinical practice known to one of ordinary skill in the art. In a particular embodiment, the pharmaceutical composition is contained in an implant that is suitable for implantation in the mammalian recipient. Thus, the methods and compositions of the invention provide for an implantable bacteriophage gene therapy vector that is useful for delivering a therapeutic polynucleotide to the mammalian recipient over an extended period of time.

According to another aspect of the invention, a kit is provided which contains: (1) a first container containing an encapsidated bacteriophage having an appropriate surface marker and a genome (preferably, a genome having multiple cloning sites, such as the lambda DASH II genome) and instructions for inserting exogenous genetic material, which may contain coding sequence and upstream and downstream regulatory elements into the genome, preferable into the multiple cloning site, (2) a second container containing an agent for attaching a ligand to the surface of the bacteriophage, wherein the ligand is designed to bind to a receptor on the surface of a mammalian cell; and (3) instructions for attaching the agent to the surface of the bacteriophage. More preferably, the kit further includes instructions for transducing a desired target mammalian cell. Alternatively, the desired ligand coding sequence is included in the genome of the bacteriophage and the agent for attaching the ligand to the surface is unnecessary.

These and other aspects of the invention, as well as various advantages and utilities, will be more apparent with reference to the detailed description of the preferred embodiments and to the accompanying drawings.

All references, patent publications and patents identified in this disclosure are incorporated in their entirety herein by reference.

Brief Description of the Drawings

Fig. 1, including 1A, 1B and 1C, is a schematic diagram describing the experimental strategy for generation of targeted bacteriophage vectors by chemical modification and use of the modified bacteriophage for gene delivery to cells of specific mammalian tissue origin.

Fig. 2 is a schematic diagram describing the experimental strategy for generation of targeted bacteriophage vectors by fusing the coding DNA sequences of a ligand "L" into the bacteriophage virion capsid specific "D" gene.

Fig. 3 is a schematic representation of a chimera of lambda DASH II /CMV promoter enhancer/beta-galactosidase gene sequences. Fig. 3A shows a restriction enzyme map and schematic representation of a restriction digest of the bacteriophage vector. Fig. 3B shows a schematic representation of the CMV promoter enhancer/beta-galactosidase gene sequences and ligation of this DNA to the digested bacteriophage to form the chimera, followed by (a) in vitro packaging, (b) propagation in E. coli, and (c) purification by CsCl gradient centrifugation to form a bacteriophage that is capable of expressing the beta-galactosidase (beta-gal) gene under the control of the CMV promoter in mammalian cells. The beta-gal gene contains a nuclear localization signal that directs localization of the translation product to the nucleus.

Fig. 4 is a schematic representation of a chimera of lambda DASH II/PGK promoter/-galactosidase gene sequences that are formed as described above in Fig. 3.

Fig. 5 is a schematic representation of a MCK/DMD/lambda bacteriophage chimeric DNA construct.

Fig. 6 is a schematic representation of a heregulin/lambda bacteriophage chimeric DNA construct in which a portion of the heregulin cDNA is fused, in frame, with the 3' end of the wild type capsid D-gene. Fig. 6A shows generation of the polynucleotide fragments; Fig. 6B shows the joining of the polynucleotide fragments; and Fig. 6C shows the generation of a modified bacteriophage expressing the heregulin-protein D chimeric capsid genes.

Fig. 7 is a schematic representation of A) the generation of targeted bacteriophage vectors by fusing a cyclic RGD ligand onto the bacteriophage lambda virion head specific D-gene product; B) generation of fragments for gene fusion; C) joining the fragments; D) generation of targeted lambda phage vector expressing the cyclic RGD-D chimeric capsid and containing the CMV beta-gal reporter gene; E) generation of cyc RGD modified lambda DASH II bacteriophage containing the murine dystophin gene expression cassette; and F) generation of cyc RGD modified lambda DASH II bacteriophage containing the Factor VIII/IRES/Von

Willebrand's Factor gene expression cassette. In reference to fig. 7B, the primers are defined as follows: primer "a" (SEQ. ID NO. 18) contains only wt sequence of the lambda DASH II "C" gene; primer "b" (SEQ. ID NO. 19) has a 3' end that is complementary to the 3' end of the wt "D" gene and a 5' end which contains the coding sequence to cyclic RGD; primer "c" (SEQ. ID NO. 20) has a 5' end that is complementary to the 5' end of cyclic RGD and a 3' end which is complementary to the 5' end of the wt "E" gene; and primer "d" (SEQ. ID NO. 21) contains only wt sequence of the lambda DASH II "E" gene. In reference to fig. 7D, the preparation of the lambda DASH II/CMV beta-gal vector was by first digesting the vector with BsrGI and Eco R1 (*lambda coordinates) and cleaving into fragments. Fragment #2 (5220-6142bp) was then separated and removed by gel electrophoresis. Co-digestion with Eco R1 favors recombination of productive phage genome. (Fragment 1: *1-5220; Fragment 2: *5221-*6142; Fragment 3: *6143-*15855; Fragment 4: *15856-CMV-betagal-*41900; and Fragment 5: *41900-48000).

Detailed Description of the Invention

The instant disclosure provides bacteriophages for delivering an exogenous polynucleotide into a target cell, preferably a mammalian cell. A "bacteriophage", for the purposes of this invention, refers to a bacteriophage that: (1) contains exogenous genetic material that can be transcribed and, optionally, translated in a mammalian cell and (2) contains on its surface a ligand that selectively binds to a receptor on the surface of a target cell, such as a mammalian cell. As used herein, "exogenous genetic material" refers to a polynucleotide (e.g., nucleic acid or oligonucleotide), either natural or synthetic, that is not naturally found in a bacteriophage, or if it is naturally found in the bacteriophage, it is not transcribed or expressed at biologically significant levels by the bacteriophage. "Exogenous genetic material" includes a non-naturally occurring polynucleotide that can be transcribed into an anti-sense RNA, as well as all or part of a "heterologous gene" (i.e., a gene encoding a protein which is not expressed or is expressed at biologically insignificant levels in a naturally-occurring bacteriophage). Thus, for example, the instant invention embraces the introduction into a mammalian cell of an expression cassette including a recombinant gene containing an inducible promoter operably coupled to a coding sequence of a therapeutic polynucleotide. In the preferred embodiments, the exogenous genetic material of the bacteriophage can be both transcribed and translated in the mammalian target cell. Exemplary bacteriophages that satisfy at least the first of these criteria are provided in Table 1. Exemplary ligands that can be attached to the bacteriophage surface, e.g., covalently coupled to the surface, expressed, or specifically adsorbed or affinity bound thereto, are provided in Table 6. (Tables 1-8 are presented at the end of the detailed description of the invention, immediately preceding the specific Examples section.)

The bacteriophages are useful for delivering an exogenous polynucleotide into a mammalian target cell for *ex vivo* and *in vivo* gene therapy, as well as for producing exogenous polynucleotide products (e.g., transcription products such as antisense mRNA or catalytic RNAs and translation products) in culture or *in vivo*. The bacteriophages of the invention are particularly useful for delivering an exogenous polynucleotide containing between about one and one-hundred kilobases to a mammalian target cell, depending on the particular bacteriophage that is selected. For example, lambda and p1 can be used to deliver exogenous polynucleotides containing up to 9 to 23 kb and up to 75 to 100 kb, respectively. A minimal lambda cassette can deliver exogenous polynucleotides containing up to 50 kb. In contrast, conventional viral vectors for gene therapy viral vectors can accommodate, at best, a polynucleotide containing up to about

seven kilobases for delivery to a mammalian cell. Thus, the instant invention advantageously provides a method for delivering relatively large genes and/or multi-gene complexes to a mammalian cell for gene therapy purposes and for the *in vitro* or *in vivo* production of gene products. In general, the bacteriophages of the invention can be propagated in prokaryotic cells. Accordingly, the cost of producing the bacteriophages of the invention is relatively inexpensive compared to the cost of producing more conventional gene therapy vectors, such as retroviruses, adenovirus, or adeno-associated virus.

According to one aspect of the invention, an improved method for gene therapy is provided. The improved gene therapy method utilizes a bacteriophage as a vector to introduce a therapeutic polynucleotide into a target cell of a mammalian recipient. The improved gene therapy method involves two steps: (1) contacting the bacteriophage with the target cell under conditions (a) to permit selective binding of a ligand on the surface of the bacteriophage to a receptor on the surface of the target cell and (b) to allow the bacteriophage to enter the target cell; and (2) allowing the target cell to live under conditions such that the therapeutic polynucleotide is transcribed therein. The mammalian recipient is diagnosed as having a medical condition that is treatable by administration to the recipient of the therapeutic polynucleotide or a product thereof. In the preferred embodiments, the mammalian recipient is a human. Exemplary medical conditions and their respective therapeutic polynucleotides (or products thereof) that are useful for treating these conditions are provided in Table 2. In general, the medical conditions that are treatable in accordance with the methods of the invention include genetic diseases (i.e., diseases that are attributable to one or more gene defects) and acquired pathologies (i.e., pathological condition that are not attributable to an inborn genetic defect). The improved method for gene therapy also embraces prophylactic processes (i.e., delaying the onset of the foregoing medical conditions).

The bacteriophage genome contains a therapeutic polynucleotide that encodes a therapeutic polynucleotide product which is useful for treating (i.e., delaying the onset, inhibiting or reducing the symptoms of) the medical condition. As used herein, a "therapeutic polynucleotide" refers to a polynucleotide that mediates a therapeutic benefit in a recipient of the polynucleotide or product thereof. A therapeutic benefit may be an alteration of cell proliferation, a change of expression of a single or multiple genes or proteins, a cytotoxic effect against a pathogen, inhibition of viral replication, replacement of a defective gene and the like. Therapeutic polynucleotides may be administered in the form of a polynucleotide operably

joined to regulatory sequences, disposed in the bacteriophage vector for replication or regulated expression, or in separate non-operable pieces that can become operably joined in the target cell to yield an operable expression system. Therapeutic polynucleotides include genes encoding the transcription and translation products identified in Table 2. Therapeutic polynucleotides also embrace polynucleotides that encode diagnostic agents that can be detected *in situ* or *ex vivo* and that are useful in diagnosing a medical condition. Therapeutic polynucleotides that encode diagnostic agents include the genes encoding, for example, an enzyme that catalyzes a reaction, *in situ*, to yield a detectable product. Thus, as used herein, a "therapeutic polynucleotide product" refers to a molecule produced as a result of transcription or translation of the therapeutic polynucleotide. Therapeutic polynucleotide products include transcription products (e.g., antisense mRNA and catalytic RNA) and translation products (e.g., proteins or peptides) of the therapeutic polynucleotide.

Antisense oligonucleotides that have been approved for gene therapy protocols and/or clinical trials are provided in Table 2. As used herein, the phrases "antisense oligonucleotides" or "antisense" describe an oligoribonucleotide which hybridizes under physiological conditions to DNA comprising a particular gene or to an RNA transcript of that gene and, thereby, inhibits the transcription of that gene and/or the translation of the mRNA. The antisense molecules are designed so as to hybridize with the target gene or target gene product and thereby, interfere with transcription or translation of the target mammalian cell gene. Those skilled in the art will recognize that the exact length of the antisense oligonucleotide and its degree of complementarity with its target will depend upon the specific target selected, including the sequence of the target and the particular bases which comprise that sequence. It is preferred that the antisense oligonucleotide be constructed and arranged so as to bind selectively with the target under physiological conditions, i.e., to hybridize substantially more to the target sequence than to any other sequence in the target cell under physiological conditions. Based upon the known sequence of a gene that is targeted for inhibition by antisense hybridization, or upon allelic or homologous genomic and/or cDNA sequences, one of skill in the art can easily choose and synthesize any of a number of appropriate antisense molecules for use in accordance with the present invention. In order to be sufficiently selective and potent for inhibition, such antisense oligonucleotides should comprise at least 7 and, more preferably, at least 15 consecutive bases which are complementary to the target. Most preferably, the antisense oligonucleotides comprise a complementary sequence of 20-30 bases. Although oligonucleotides may be chosen which are

antisense to any region of the gene or RNA (e.g., mRNA) transcripts, in preferred embodiments the antisense oligonucleotides are complementary to 5' sites, such as translation initiation, transcription initiation or promoter sites, that are upstream of the gene that is targeted for inhibition by the antisense oligonucleotides. In addition, 3'-untranslated regions or telomerase binding sites may be targeted. Furthermore, 5' or 3' enhancers may be targeted. Targeting to mRNA splice sites has also been used in the art. In at least some embodiments, the antisense is targeted, preferably, to sites in which mRNA secondary structure is not expected (see, e.g., Sainio et al., Cell Mol. Neurobiol. 14(5):439-457 (1994)) and at which proteins are not expected to bind. The bacteriophages of the invention, by virtue of their ability to accommodate therapeutic polynucleotides which are relatively large in size, are particularly useful for delivering to the target cell a polynucleotide that encodes multiple copies of the same or different mRNAs in tandem, thereby increasing the effective concentration of antisense mRNA in the target cell.

The selective binding of the antisense oligonucleotide to a mammalian target cell nucleic acid effectively decreases or eliminates the transcription or translation of the mammalian target cell nucleic acid molecule. Reduction in transcription or translation of the nucleic acid molecule is desirable in those medical conditions where transcription and translation of the mammalian target cell nucleic acid leads to an adverse medical condition. For example, the antisense oligonucleotides of the invention can be used to reduce the expression of oncogenes to treat cancers whose proliferation is mediated by expression of these oncogenes.

The bacteriophages of the invention are also useful for delivering therapeutic polynucleotides that encode specific antigen peptides to antigen presenting cells for processing and presentation at the cell surface to enhance the immune system response of the mammalian recipient to a specific peptide antigen. Exemplary peptide antigens that can be expressed to induce or otherwise enhance an immune response are shown in Table 7. In the preferred embodiments, the therapeutic polynucleotide encodes one or more peptide antigens that vaccinate the mammalian recipient against a tumor, a virus, a bacteria, or a parasite. Optionally, auxiliary therapeutic polynucleotides are inserted into the bacteriophage genome to enhance or otherwise improve the therapeutic efficacy of the therapeutic polynucleotide product in treating the condition. Exemplary auxiliary polynucleotides for delivery to the mammalian target cell include polynucleotides encoding tumor suppressor genes, polynucleotides encoding antisense mRNA or encoding catalytic RNA that inactivate oncogenes, and polynucleotides that render a

target tumor cell more susceptible to an administered drug (e.g., suicide genes encoding, for example, thymidine kinase). Auxiliary polynucleotides also include polynucleotides encoding cytokines that enhance a naturally occurring anti-tumor immunity. Exemplary cytokines which have this function include, e.g., IL-4, TNF, IL-2, and GM-CSF.

The therapeutic polynucleotide is inserted into the bacteriophage genome using conventional recombinant DNA techniques. See, e.g., Methods in Enzymology, vol. 152, ed. S. L. Berger, A.R. Kimmel (1987) Academic Press, New York, NY. In the preferred embodiments, the bacteriophage is a lambda phage and the therapeutic polynucleotide is inserted into well-defined restrictions sites in the lambda phage. (See, e.g., the Examples and figures.) Optionally, recombination sequences (i.e., polynucleotides having a nucleic acid that allows homologous recombination) are provided at the 5' and 3' ends of the therapeutic polynucleotide to permit site-directed insertion of the therapeutic polynucleotide into a preselected location in the genomic DNA of the target cell via homologous recombination.

In contrast to the viral vectors that presently are available for human gene therapy, the bacteriophages of the invention can accommodate a therapeutic polynucleotide containing between about one and up to one-hundred kilobases, depending upon the particular bacteriophage selected. For example, lambda bacteriophages, e.g. lambda DASH II, and p1 phage can accommodate up to about 9 to 23 kb and up to about 75 to 100 kb, respectively. Minimal lambda cassette, such as described below, can accommodate up to about 50 kb. Preferably, the therapeutic polynucleotide contains between about 10 and 90 kilobases, more preferably, the therapeutic polynucleotide contains between about 15 and 85 kilobases. Thus, the improved gene therapy method disclosed herein is particularly useful for gene therapy applications which require administration of a single therapeutic polynucleotide (or a product thereof) having a size within the foregoing kilobase range, as well as for delivering multiple therapeutic polynucleotides which, together, have a size within this kilobase range. Exemplary therapeutic polynucleotides containing more than 7 kilobases include dystrophin, members of the globin gene complex, clotting factor VIII, von Willebrand's factor, collagen type VII, fibrillin, and any other gene(s)/gene complexes than are too large to deliver (efficiently) to mammalian cells using conventional viral vectors. Additional therapeutic polynucleotides that can be delivered in accordance with the methods of the invention and that fall within the preferred kilobase size ranges can be identified by, for example, referring to the GenBank or other gene sequence data bases. See, also, Table 3 for a list of preferred therapeutic polynucleotides that can

be delivered using the bacteriophages disclosed herein. The delivery of a therapeutic polynucleotide containing more than 7 kilobases has not been possible using conventional gene therapy viral vectors. A summary of the insert size limitations for conventional gene therapy vectors compared to the bacteriophages disclosed herein is provided in Table 8.

The packaging and engineering of the lambda bacteriophage vector can be modified to permit the vector to accommodate up to approximately 50 kb of exogenous coding sequence. This approach involves engineering a recombinant cosmid vector DNA construct that contains an antibiotic resistance gene, e.g. ampicillin, a lambda origin of replication, and a DNA insert up to 50 kb in size flanked by COS (CoheSive ends of wild type bacteriophage lambda genome) sites. This cosmid can be replicated to very high copy numbers in standard strains of E. Coli and then can be isolated using standard techniques for use in the packaging as described below.

The second component of this modified packaging system is a COS-negative lambda lysogen strain of bacteria in whose bacterial chromosome is integrated the structural proteins and enzymes requisite for packaging of an infective lambda virus. Using standard recombinant techniques (Molecular Cloning, 2nd Edition, Sambrook et al., Cold Spring Harbor Laboratory, 1989), the lysogen strain is engineered to contain modifications of certain packaging proteins such that the final modified lambda phage vector is able to effectively target the intended cell type(s), sub-cellular compartments, or sub-cellular organelles. For example, E. Coli can be infected with a modified bacteriophage that contains a fusion D gene-RGD construct Using standard methods lysogen that contain the D gene-RGD fusion construct can be selected . Combining protein extracts from a D gene minus lysogen with a D gene-RGD fusion lysogen will provide the full complement of necessary packaging proteins. Using this combination of lysogen extracts, bacteriophage genomes or minimal lambda cassettes can be encapsidated in vitro with a D gene-RGD fusion protein that can target the recombinant virus to cells expressing RGD's cognate receptor. Preferably, the D-gene is modified to include in frame a ligand, e.g. cyclic RGD, so that the vector is internalized by the targeted cell type(s). This lysogen strain is incapable of producing virus because the COS sites are absent. The lysogen strains are grown to large quantities using standard bacterial culture techniques. Then, the lambda packaging proteins, including any modified forms of these proteins, can be obtained by standard methods, such as freeze thawing and sonication of lysogen.

The packaging of the high capacity bacteriophage lambda vector is accomplished by mixing the purified engineered cosmid DNA with the isolated protein extract from the above

lysogen strain at approximately room temperature. This mixing results in the packaging of replication deficient modified bacteriophage lambda virus particles that contain both surface proteins as determined by the modified lysogen strain from which the packaging proteins are derived and a genome of an insert of up to about 50 kb flanked by COS sites yet lacking other bacteriophage coding sequences.

In the preferred embodiments, the bacteriophage genome further includes a regulatory sequence, e.g., a promoter region (also referred to as a "promoter"), that is operably coupled to the therapeutic polynucleotide. The regulatory sequence controls the expression of the therapeutic polynucleotide in the target cell. As used herein, a therapeutic polynucleotide (also referred to as "coding sequence") that encodes a therapeutic polynucleotide product , and regulatory sequences are said to be "operably" joined when they are covalently linked in such a way as to place the transcription or the expression of the coding sequence under the influence or control of the regulatory sequences. If it is desired that the coding sequences be translated into a functional protein, two DNA sequences are said to be operably joined if induction of a promoter in the 5' regulatory sequence results in the transcription of the coding sequence and if the nature of the linkage between the two DNA sequences does not (1) result in the introduction of a frame-shift mutation, (2) interfere with the ability of the promoter region to direct the transcription of the coding sequences, or (3) interfere with the ability of the corresponding RNA transcript to be translated into a protein. Thus, a promoter region would be operably joined to a coding sequence if the promoter region were capable of effecting transcription of that DNA sequence such that the resulting transcript might be translated into the desired protein or polypeptide.

The precise nature of the regulatory sequences needed for gene expression may vary between species or cell types, but shall in general include, as necessary, 3' or 5' non-transcribed and/or non-translated sequences involved with the initiation of transcription and translation respectively, such as a TATA box, CAAT sequence, and the like. Especially, such 5' non-transcribing regulatory sequences will include a promoter region which includes a promoter sequence for transcriptional control of the operably joined gene. Regulatory sequences may also include enhancer sequences or upstream 5' or downstream 3' transcriptional regulatory sequences as desired. The bacteriophages of the invention optionally include 5' leader signal, or membrane integration sequences 5' or 3'.

Exemplary promoters that are useful for protecting the instant invention are provided in Table 5, which shows both constitutive promoters and regulatable promoters (e.g., cell lineage

specific promoters, inducible promoters). Exemplary constitutive promoters also are included in Table 5. In the preferred embodiments, the constitutive promoter is selected from the group consisting of a promoter of the phosphoglycerokinase gene, a long terminal repeat (LTR) of retroviruses, e.g., Rous sarcoma virus, Moloney murine leukemia virus. Exemplary tissue or cell specific transcriptional regulatory sequences are derived from the genes encoding the following proteins: tyrosinase, lipoprotein lipase, albumin, muscle creatine kinase, keratin (K14/K10), globin gene cluster, immunoglobulin heavy chain gene cluster, and involucrin. Several liver-specific promoters, such as the albumin promoter/enhancer, also have been described (see, e.g., PCT application number PCT/US95/11456, having international publication number WO96/09074, entitled "Use of a Non-mammalian DNA Virus to Express an Exogenous Gene in a Mammalian Cell," hereinafter W0 96/09074, and the references cited therein). In particular, the alpha-feto protein promoter, can be used to effect expression of a therapeutic polynucleotide(s) in liver tumor cells (but not normal liver cells) for treating liver cancer. Exemplary inducible promoters are identified in Table 5 and are described in the following references: Science 268:1786 (1995); TIBS 18:471 (1993); PNAS 91:3180 (1994); PNAS 90:1657 (1993); PNAS 88:698 (1991); Nature Biotechnol. 14:486 (1996); and PNAS 93:5185 (1996). The preferred inducible promoter system is the tetracycline inducible system. An exemplary repressible promoter, the tetracycline repressible system, is identified in Table 5 and is described in PNAS 89:5547 (1992).

A tetracycline inducible promoter system includes, for example, two tandem constructs: an appropriate promoter operably linked to a trans-activating polypeptide coding sequence (rtTA) (a mutated Tet R linked to a VP16) and poly A signal in tandem with a tetracycline responsive element (tetO and a eukaryotic minimal promoter) operably coupled to a coding sequence of interest. Containing a poly A site, these two tandem constructs can be, optionally, joined into a single construct separated by an internal ribosomal entry site (IRES). The tetO driven coding sequence is 5' to the mammalian promoter driven coding sequence. Addition of tetracycline, doxycycline, or derivatives thereof, activates the transactivating polypeptide to bind tetO and, in turn, to drive transcription of a polycistronic message including first the coding sequence of interest and the transactivator. The IRES allows initiation of translation of the trans-activating polypeptide or the coding sequence of interest, independently, from a single polycistronic message. In the absence of tetracycline, the transactivator is not activated and the tetO driven transcription of the coding sequence of interest is substantially reduced to negligible

levels.

Preferably, the bacteriophage genome further includes an enhancer region ("enhancer"). Exemplary enhancers that are useful for practicing the instant invention are provided in Table 5. The preferred enhances are selected from the group consisting of the following: a locus control region (beta-globin enhancer), an immunoglobulin gene enhancer, a cytomegalovirus (CMV) enhancer, a muscle creatine kinase enhancer, and an SV40 enhancer. Optionally, the bacteriophage genome can be engineered to contain an origin of replication to effect autonomous replication and facilitate persistence of the therapeutic polynucleotide in the mammalian cell. Origins of replication derived from mammalian target cells have been identified (see, e.g., Burhans, et al., 1994, Science 263: 639-640).

The bacteriophages optionally contain one or more sequences that are suitable for use in the identification of cells that have or have not been transfected. "Transfection", as used herein, refers to the introduction of the bacteriophage genome into the target cell. Markers to identify cells that have been transfected include, for example, genes encoding proteins that increase or decrease resistance or sensitivity to antibiotics or other compounds, genes which encode enzymes having activities that are detectable by standard assays known in the art and genes which detectably (e.g. visibly) effect the phenotype of the transfected target cells, hosts, or plaques. Exemplary genes that are suitable as markers include the lacZ genes, chloramphenicol acetyltransferase genes, alkaline phosphatase genes, luciferase genes, and green fluorescent protein genes.

The bacteriophages of the invention contain on their surface a ligand that selectively binds to a receptor expressed on the surface of the target cell. The bacteriophage ligand binds to the target cell receptor to form a ligand-receptor complex that is internalized by the target cell. In this manner, the bacteriophage can be targeted for delivery to a pre-selected cell or tissue type, i.e., by selecting a ligand/receptor pair wherein the receptor is selectively expressed on a specific population of cells. Exemplary pairs of ligands/receptors which are useful in accordance with the methods of the invention are provided in Table 6.

The preferred ligand/receptor pairs for use in accordance with the methods of the invention include the following: insulin/insulin receptor, heregulin/heregulin receptor, keratinocyte growth factor/keratinocyte growth factor receptor, hepatic growth factor/hepatic growth factor receptor, RGD peptides/integrin alpha-5:beta-1, interleukin-2/interleukin-2 receptor, galactose/asialoglycoprotein, low density lipoprotein (LDL) or apoB100/LDL receptor,

very low density lipoprotein (VLDL), apoE/VLDL receptor, or HDL or apoA1/HDL receptor. In addition to these ligand/receptor pairs, antibodies can be attached to the bacteriophage via interaction of a bacteriophage coat protein specific antibody with its cognate antigen or via interaction of a recombinant bacteriophage coat protein that contains avidin and a biotinylated antibody. Once the antibody is tightly bound to the bacteriophage as described above, the antibody can direct the bacteriophage either to cells that express the Fc receptor in the case of a coat protein specific IgG antibody or to cells that express the cognate antigen of the attached biotinylated antibody. Selective ligand-receptor interaction also is useful for mediating internalization of the bacteriophage into subcellular locations e.g., the nucleus, mitochondria, and other membranes-bound organelles or cytoplasmic molecular aggregates of protein and/or nucleic acid.

In addition to the well-known ligand/receptor pairs for delivering a ligand-labeled component to a particular cell type, novel ligands can be identified using phage display procedures such as those described in (S. Hart, et al., J. Biol. Chem. 269(17):12468 (1994)). While such filamentous phages could, of course, never be used to deliver genetic material to a cell (because they are single stranded), this methodology is potentially very useful in the discovery of novel receptor ligand interactions. In general, phage display libraries using, e.g., M13 or fd phage are prepared using conventional procedures such as those described in the foregoing reference. The libraries display inserts containing from 4 to 80 amino acid residues. The inserts optionally represent a completely degenerate or a biased array of peptides. Ligands that bind selectively to a particularly type of target cell (e.g., mammalian cell) are obtained by selecting those phages which express on their surface a ligand that binds to the target cell of interest. These phages then are subjected to several cycles of reselection to identify the peptide ligand-expressing phages that have the most useful binding characteristics. Typically, phages that exhibit the best binding characteristics (e.g., highest affinity) are further characterized by nucleic acid analysis to identify the particular amino acid sequences of the peptides expressed on the phage surface and the optimum length of the expressed peptide to achieve optimum binding to the target mammalian cell. Alternatively, such peptide ligands can be selected from combinatorial libraries of peptides containing one or more amino acids. Such libraries can further be synthesized which contain non-peptide synthetic moieties which are less subject to enzymatic degradation compared to their naturally-occurring counterparts. These novel ligands can be attached to the bacteriophage surface to deliver the phage to the particular target cell of

interest. Alternatively, the ligands may be selected from polynucleotide libraries, as polynucleotides can also demonstrate specific affinity for polypeptide cell surface receptors. Exemplary screening methods for selecting bacteriophage vectors that transduce mammalian target cells via receptor-mediated endocytosis and target cells that internalize the bacteriophage vectors by this mechanism are described in the Examples.

At least four general procedures are available for attaching a ligand to the surface of a bacteriophage. These include (1) chemical modification of the bacteriophage surface (e.g., galactosylation, cross linking reactions); (2) modification of the bacteriophage genome to express a ligand on the bacteriophage surface (e.g., a fusion protein formed between the ligand and a functional viral packaging protein); (3) selective binding of a ligand (e.g., a monoclonal antibody, a polyclonal antibody, or functionally active fragments thereof containing an Fc domain) to a bacteriophage surface antigen to mediate targeting of the bacteriophage to cells that express an Fc receptor on their surface; and (4) modification of the bacteriophage genome to express a surface avidin-bacteriophage coat protein fusion product to which a biotinylated ligand (e.g., antibody) can be attached.

The simplest method for attaching a ligand to the surface of a bacteriophage is a chemical modification reaction in which the surface of the bacteriophage is subjected to galactosylation or lactosylation via N-linked glycosidic covalent linkages so as to attach galactose or lactose, respectively, to the bacteriophage surface. See, e.g., the Examples. Galactose- or lactose-labeled bacteriophages selectively bind to asialoglycoprotein receptors on the surface of hepatocytes to form a ligand-bacteriophage complex that is internalized by the target cell. Chemical modification also can be used to attach a peptide ligand to the bacteriophage surface. For example, peptide ligands containing a free amine group, carboxyl group, or sulfhydryl group can be attached to the bacteriophage surface using conventional procedures known to those of ordinary skill in the art for cross linking proteins. See, e.g., U.S. Patent No. 5,108,921, issued to Low et al. If the peptide ligand does not have a free amine or carboxyl group, such a group can be introduced by, for example, introducing a cysteine (containing a reactive thiol group) into the peptide ligand by site directed mutagenesis. Disulfide linkages can be formed between thiol groups in, for example, the peptide ligand and a protein expressed on the surface of the bacteriophage. For example, covalent linkages can be formed using bifunctional crosslinking agents that are known by those of ordinary skill in the art to have utility with respect to crosslinking peptides and proteins. Exemplary crosslinking agents include bismaleimidohexane

(which contains thiol-reactive maleimide groups and which forms covalent bonds with free thiols). See, also, the Pierce Co. Immunotechnology Catalogue and Handbook Vol. 1 for a list of exemplary homo- and hetero-bifunctional crosslinking agents, thiol-containing amines and other molecules with reactive groups for a comprehensive list of commercially available agents and corresponding peptide coupling chemistries that can be used to attach a peptide ligand to, for example, an amino acid functional group (e.g., amine) on the surface of a bacteriophage. Further exemplary coupling chemistries that are suitable for this purpose include methods which utilize the following crosslinking agents: glutaraldehyde (M. Riechlin, Methods in Enzymology 70:159-165 (1980); N-ethyl-N'-(3-dimethylaminopropyl)-carbodiimide (T.L. Goodfriend, et al., Science 144:1344-1346 (1964); and N-ethyl-N'-(3-dimethylaminopropyl)-carbodiimide (M.H. Klapper and I.M. Klotz, Methods in Enzymology 25:531-536 (1972)). In general, the ligand-labeled bacteriophages can be prepared by using well-known methods for forming amide, ester or imino bonds between acid, aldehyde, hydroxy, amino, or hydrazo groups on the respective bacteriophage and ligand. In certain embodiments, reactive functional groups that are present in the amino acid side chains of the bacteriophage extracellular proteins or peptide ligands preferably are protected with a protecting group prior to coupling the ligand to the bacteriophage. As used herein, "protecting group" refers to a molecule which is bound to a functional group on a bacteriophage or peptide ligand and which may be selectively removed therefrom to expose the functional group in a reactive form. Preferably, the protecting groups are reversibly attached to the functional groups and can be removed therefrom using, for example, chemical or other cleavage methods. Thus, commercially available side-chain-blocked amino acids (e.g., FMOC-derived amino acids from Advanced Chemtech. Inc., Louisville, KY) can be used to synthesize peptide ligands in accordance with standard peptide synthesis methods to obtain peptide ligands having side-chain-blocked amino acids. Alternatively, the peptide ligand side chains can be reacted with protecting groups after peptide ligand synthesis, but prior to the covalent coupling reaction. In this manner, the ligand-labeled bacteriophages of the invention can be prepared in which the amino acid side chains of the peptide ligand do not participate to any significant extent in the coupling of the peptide ligand to functional groups on the surface of the bacteriophage.

Alternatively, genetic engineering methods can be used to attach a ligand to the surface of a bacteriophage. For example, a sequence encoding the ligand ("ligand coding sequence") can be inserted, in frame, into the bacteriophage genome adjacent to or within a bacteriophage protein that is expressed on the bacteriophage surface to provide a fusion protein that contains

both functional ligand and a functional viral packaging protein. (See, e.g., the Examples.)

According to yet another embodiment, the ligand-labeled bacteriophage is formed by selectively binding an antibody or functionally active fragment thereof (i.e., an antibody fragment containing at least one antigen-binding site) to an antigen that is contained on the surface of the bacteriophage (e.g., a bacterial extracellular protein). The selectively bound antibodies mediate targeting of the bacteriophage to a target cell that contains on its surface an Fc receptor (e.g., a phagocyte or antigen presenting cell). Antibodies that are useful in accordance with this aspect of the invention are antibodies that exhibit a sufficiently high binding affinity for a bacteriophage antigen to result in little or no dissociation of the antibody-antigen complex under physiological conditions. In general, such antibody binding to the surface of the bacteriophage is performed by contacting the bacteriophage with an antibody that selectively binds to an antigen expressed on the bacteriophage surface under the same conditions that are used for performing an immunoassay, e.g., an ELISA, RIA. Exemplary conditions are described in Current Protocols in Immunology, ed. Coligan, J.E., et al., National Institutes of Health, John Wiley and Sons, Inc. (1994).

In a particularly preferred embodiment, the ligand is attached to the surface of the bacteriophage by means of an avidin/biotin complex. As used herein, "avidin" or "avidin peptide" refers to an avidin molecule, a streptavidin molecule, or a fragment or variant thereof that binds to biotin with an affinity that is approximately the same (i.e., within 10%) or greater than the affinity with which streptavidin binds to biotin. According to this embodiment, the bacteriophage is modified to express on its surface avidin or a portion thereof that selectively binds to biotin with the requisite binding affinity. Modification of the bacteriophage to express avidin is most easily accomplished by inserting the nucleic acid encoding avidin or a functionally active portion thereof into the bacteriophage genome such that the avidin or avidin portion is expressed on the bacteriophage surface. For example, the avidin can be inserted, in frame, into the D gene of the lambda coat protein, using well-defined restriction sites in the lambda phage. (See, e.g., the Examples and figures.) In this manner, an avidin-expressing bacteriophage is produced which serves as an intermediate for attachment of a biotinylated ligand to the bacteriophage surface.

Alternatively, avidin or a functionally active portion thereof can be chemically coupled to the bacteriophage surface using standard cross-linking chemistries, such as those described above. The avidin-labeled bacteriophage permits non-covalent, yet high affinity, attachment of

pre-selected biotinylated ligands to the bacteriophage surface for receptor-mediated targeted delivery to the mammalian target cell. Exemplary ligands which can be biotinylated in accordance with standard procedures are provided in Table 6. Alternatively, the bacteriophage can be biotinylated and an avidin-labeled ligand can be used to form the ligand-labeled bacteriophages described herein.

The bacteriophages of the invention are contacted with the target cell under conditions to permit selective binding of the ligand on the surface of the bacteriophage to the receptor on the surface of the target cell and to allow the bacteriophage to enter the target cell. Conditions which permit the binding of a receptor to its cognate ligand are the physiological conditions (e.g., the particular pH, ionic strength, viscosity) at which the ligands and receptors are known to bind to one another in vivo and the conditions at which the ligands and receptors are known to bind to one another in vitro, such as in receptor assays for determining the presence of a ligand in, for example, a biological fluid. Such conditions are known to those of ordinary skill in the art of receptor-mediated processes, such as receptor-based binding assays and receptor-mediated delivery of therapeutic agents to preselected tissues *in situ*.

In general, the conditions that allow the target cell to live and transcribe the therapeutic polynucleotide are the same conditions that permit selective binding of the ligand to the receptor and that allow the bacteriophage to enter the target cell. Optionally, the conditions that allow the cell to transcribe the therapeutic polynucleotide further include the addition of an inducer (see, e.g., Table 5) that activates an inducible promoter to induce transcription and translation of the therapeutic polynucleotide. The optimum conditions for inducing the transcription and translation of a therapeutic polynucleotide that is under the control of a particular inducible promoter can be determined by one of ordinary skill in the art using no more than routine experimentation. In general, for *in vitro* gene therapy, conventional tissue culture conditions and methods are used to sustain the mammalian cell in culture. For example, the mammalian cell can be allowed to live on a substrate containing collagen, e.g., type I collagen, or a matrix containing laminin, such as described in PCT application number PCT/US95/11456, having international publication number WO96/09074, entitled "Use of a Non-mammalian DNA Virus to Express an Exogenous Gene in a Mammalian Cell," and the references cited therein.

As used herein, "contacting", in reference to the bacteriophage and the target cell, refers to bringing the bacteriophage into sufficiently close proximity to the target cell to permit the receptor on the target cell to selectively bind to the ligand on the bacteriophage. Such conditions

are well known to those of ordinary skill in the art and are exemplified by the procedure provided in the Examples. See also, e.g., U.S. patent No. 5,108,921, issued to Low et al. which reports the conditions for receptor-mediated delivery of "exogenous molecules" such as peptides, proteins and nucleic acids to animal cells and U.S. patent No. 5,166,320, issued to Wu et al., which reports the conditions for the receptor mediated delivery of a ligand-gene conjugate to a mammalian cell. For a further discussion of the conditions and mechanisms by which receptor mediated delivery can be used to deliver an exogenous molecule into a target cell, and in particular, into a mammalian cell, see, e.g., S. Michael, et al., J.Biol., Chem. 268(10):6866 (1993), "Binding-incompetent Adenovirus Facilitates Molecular Conjugate-mediated Gene Transfer by the Receptor-mediated Endocytosis Pathway"; M. Barry, et al., Nature Medicine 2(3):299 (1996), "Toward cell-targeting gene therapy vectors: Selection of cell-binding peptides from random peptide-presenting phage libraries"; S. I. Michael, Gene Ther. 2:660 (1995), "Addition of a short peptide ligand to the adenovirus fiber protein".

The bacteriophage can be contacted with the targeted mammalian cell *in vitro*, for example, for *ex vivo* gene therapy or production of a catalytic RNA or recombinant protein in cell culture, or *in vivo* for *in vivo* gene therapy or *in vivo* production of a polynucleotide transcription or translation product. As used herein, a "mammalian target cell" refers to a mammalian cell (preferably, a human cell) which contains on its surface a receptor for the ligand that is contained (e.g., expressed) on the surface of the bacteriophage. Essentially, any mammalian cell can be targeted in accordance with the methods of the invention. The cell may be a primary cell or may be a cell of an established cell line. Exemplary cell types that can be targeted in accordance with the methods of the invention are provided in Table 4. Preferably, the mammalian cell is a hepatocyte (liver cell), a breast epithelial cell, a keratinocyte, a melanocyte, or a hematopoietic cell, e.g., erythrocyte, leukocyte, monocyte, or a lymphocyte. Screening methods, such as those described in the Examples, can be used to confirm that these and other target cells internalize the bacteriophage vectors of the invention via receptor-mediated endocytosis and, further, that these target cells express detectable levels of the exogenous polynucleotide insert. Such high-throughput screening methods can be used to select target cells that satisfy the above-noted criteria using no more than routine experimentation. In addition, such screening assays are predictive of receptor-mediated endocytosis of target cells in vivo.

Where the bacteriophage is contacted with the cell *in vitro*, the target cell subsequently can be introduced into the mammal (e.g., into the portal vein or into the spleen) if desired.

Accordingly, expression of the therapeutic polynucleotide is accomplished by allowing the cell to live or propagate *in vitro*, *in vivo*, or *in vitro* and *in vivo*, sequentially. Similarly, where the invention is used to express a therapeutic polynucleotide in more than one cell, a combination of *in vitro* and *in vivo* methods are used to introduce the therapeutic polynucleotide into more than one mammalian cell.

In *ex vivo* gene therapy, the cells are removed from a subject and a therapeutic polynucleotide is introduced (i.e., transfected) into the cells *in vitro*. Typically, the transfected cells are expanded in culture before being reimplanted into the mammalian recipient. The procedure for performing *ex vivo* gene therapy is outlined in U.S. Patent 5,399,346 and in exhibits submitted in the file history of that patent, all of which are publicly available documents. In general, *ex vivo* gene therapy involves the introduction *in vitro* of a functional copy of a gene into a cell(s) of a subject which contains a defective copy of the gene, and returning the genetically engineered cell(s) to the subject. The functional copy of the gene is under the operable control of regulatory elements which permit expression of the gene in the genetically engineered cell(s). In *in vivo* gene therapy, the target cells are not removed from the patient. Rather, the therapeutic polynucleotide is introduced into the cells of the mammalian recipient *in situ*, i.e., within the recipient. In general, the improved gene therapy method disclosed herein is practiced by using the ligand-labeled bacteriophage of the invention in place of the gene therapy vectors disclosed in the prior art (e.g., adenoviral vectors, modified adenoviral vectors, retroviral vectors, plasmids, liposomes) in the procedures currently used for administering these vectors (or cells containing these vectors) to the subjects. Such procedures are known to those of skill in the art of human gene therapy. (See, also, the references identified below with respect to *in vivo* gene therapy.)

A particularly preferred embodiment of the invention is illustrated in the Examples, which describes the construction of a lambda bacteriophage for use as a gene therapy vector. Briefly, a bacteriophage containing an exogenous polynucleotide (encoding beta-galactosidase) was chemically modified (galactosylated) to contain galactose residues on the bacteriophage surface. The galactose-labeled bacteriophage was internalized by HepG2 cells (a hepatoma cell line) in culture, presumably, by binding of the labeled bacteriophage to the asialoglycoprotein receptor present on the surface of the hepatocytes.

Insertion of the therapeutic polynucleotide into the target cell genome may be either transient or permanent. By "transient", it is meant that the bacteriophage genome lacks the

capacity to replicate and/or to segregate to progeny cells. For example, the therapeutic polynucleotide may be epigenetic and without the capacity to replicate and segregate to progeny cells (e.g., lacks an origin of replication, appropriate telomere and centromere structures). "Transient" insertion into the target cell also occurs, for example, when the bacteriophage is used to infect cells of limited replicative capacity, i.e., non-stem cells. "Permanent" insertion of the therapeutic polynucleotide into the target cell is accomplished by, for example, (1) infecting stem cells that produce bacteriophage-bearing progeny; or (2) including recombination sequences in the bacteriophage genome on either side of the therapeutic polynucleotide so as to promote reasonably efficient homologous recombination of the therapeutic polynucleotide into a defined sequence of the target cell genome or (3) random integration into the host cell chromosomal DNA. "Permanent" insertion also can be achieved by including in the bacteriophage an origin of replication, telomeres and centromeres to obtain a bacteriophage that autonomously replicates (i.e., an "artificial chromosome") and is capable of segregating into progeny cells. If the bacteriophage genome is autonomously replicating, it is preferred to further include in the bacteriophage genome appropriate enhancer-promoter sequences, such as those described in the aforementioned tables.

For *in vivo* gene therapy, the bacteriophage is administered to the mammalian recipient, for example, intravascularly, intraluminally (introduction of the bacteriophage into body cavities and lumens, such as the genital urinary tract, gastrointestinal tract, trachea-bronchopulmonary tree or other internal tubular structures), direct injection into a tissue (e.g., muscle, liver), topical application (e.g., eye drops or aerosol application to mucosal surfaces), or intracavitary (e.g., intraperitoneally or intrathecally (introduction into the cerebrospinal fluid). Optionally, an implantable pump or other device or implant (preferably, a bioerodible implant) to effect the sustained release of the bacteriophage can be used to facilitate delivery of the bacteriophage to the mammalian targeted cell over a pre-selected period of time (e.g., sustained release over a period of days to sustained release over a period of weeks to months). Although the ligand/receptor-mediated delivery of the bacteriophage is the predominant mechanism for targeting delivery of the bacteriophage to a particular cell type, delivery to the target cell can further be modulated by regulating the amount of bacteriophage administered to the mammalian recipient and/or by controlling the method of delivery. Thus, for example, intravascular administration of the bacteriophage to the portal vein or to the hepatic artery can be used to facilitate targeting the bacteriophage to a liver cell.

In general, the bacteriophage can be administered to the mammalian recipient using the same modes of administration that currently are used for adenovirus-mediated gene therapy in humans. Such conditions are adequate for contacting the bacteriophage and the target cell under conditions to permit selective binding of a ligand on the surface of the bacteriophage to a receptor on the surface of the target cell and to allow the bacteriophage to enter the target cell. These conditions are described in the following references: PNAS 90:10613 (1993); Nature Medicine 1:1148 (1995); Nature Medicine 12:266 (1996); New Engl. J. Med. 333:832 (1995); and New Engl. J. Med. 333:823 (1995). Preferably, the bacteriophage is administered to the mammalian recipient by intravascular injection, intra-organ introduction by, for example, injection into the organ or contacting the bacteriophage with the organ in the presence of a tissue permeabilizing agent; and introduction of the bacteriophage into body cavities or lumens. Optionally, immunosuppressive drugs, such as glucocorticosteroids or cyclophosphamide are co-administered with the bacteriophage to suppress a primary immune response that may be triggered by an initial exposure to a foreign antigen. Mammalian cells which have been transfected with the bacteriophage *ex vivo* can be introduced into the mammalian recipient using the known methods for implanting transfected cells into a human for gene therapy. See, e.g., U.S. Patent No. 5,399,346 ("Gene Therapy") issued to Anderson et al.; PCT International application no. PCT/US92/01890 (Publication No. WO 92/15676, "Somatic Cell Gene Therapy", claiming priority to U.S. Serial No. 667,169, filed March 8, 1991, inventor I. M. Verma); PCT International application no. PCT/US89/05575 (Publication No. WO 90/06997, "Genetically Engineered Endothelial Cells and Use Thereof", claiming priority to U.S. Serial No. 283,586, filed December 8, 1989, inventors Anderson, W.F. et al.).

The invention is not limited in utility to human gene therapy, but also can be used in the manufacture of a wide variety of proteins and nucleic acids that are useful in the fields of biology and medicine. The bacteriophages of the invention advantageously provide a method for synthesizing gene products from genes which range in size from about one to one-hundred kilobases. Further, the invention provides a method for providing the bacteriophage vectors at a low cost, namely, by propagating the bacteriophages in a prokaryotic host. Moreover, the invention provides a simple method for preparing mammalian proteins, including proper post-translational modifications, *in vitro*. Thus, the invention provides an improved method for introducing an exogenous polynucleotide into a mammalian cell. The improved method involves contacting the bacteriophages of the invention (which contain the exogenous polynucleotide)

with the mammalian cell and allowing the bacteriophage to enter the cell, gain access to the nucleus, and replicate the exogenous polynucleotide therein. As discussed above, the bacteriophage contains on its surface the ligand that selectively binds to a receptor for the ligand that is contained on the surface of the mammalian cell.

As used herein, an "exogenous polynucleotide" refers to a nucleic acid that is not normally present in the bacteriophage genome and that is inserted into the bacteriophage using recombinant engineering methodology. Examples of exogenous polynucleotides include the above described therapeutic polynucleotides, as well as regulatory polynucleotides which are not intended for therapeutic applications (e.g., polynucleotides that are introduced into the mammalian cell *in vitro* or *in vivo* for the purpose of producing a mammalian protein/protein complex *in vitro* or *in vivo)*.

According to yet another aspect of the invention, an avidin-labeled bacteriophage that is used as an intermediate in connection with the above described methods for introducing a therapeutic or other exogenous polynucleotide into a mammalian cell is provided. The intermediate can be used to prepare a bacteriophage having on its surface virtually any ligand, provided, that the ligand can be biotinylated and retain its binding activity to a receptor. Thus, the invention also provides a method for preparing a ligand-labeled bacteriophage which involves contacting an avidin-expressing bacteriophage with a biotinylated ligand under conditions to permit binding of the avidin to the biotin. As will be apparent to those of ordinary skill in the art, alternative binding pairs can be used in place of the avidin-ligand binding pairs to accomplish this same objective. Such binding pairs include, for example, streptavidin-biotin binding pairs, antibody antigen, and any other high affinity interactions. Thus, the invention provides a generic mechanism for forming a bacteriophage which contains on its surface virtually any type of ligand. The avidin-labeled bacteriophage can be provided as a component of a kit for labeling a bacteriophage with a ligand of choice. The kit includes instructions for forming a ligand-labeled bacteriophage by allowing the avidin-bacteriophage to react with a biotinylated ligand under conditions to permit selective binding of the avidin-labeled bacteriophage to the biotinylated ligand. Optionally, the kit further includes reagents, and appropriate instructions, for biotinylating a ligand of choice.

The invention provides other compositions and kits which are useful for practicing the above-described methods. According to a particularly preferred aspect of the invention, a bacteriophage of the invention is provided. The bacteriophage contains (a) a bacteriophage

genome containing an exogenous polynucleotide that can be transcribed in a mammalian cell; and (b) a ligand contained on the surface of the bacteriophage that selectively binds to a receptor expressed on the surface of a mammalian cell. The bacteriophages of the invention optionally are contained in a pharmaceutically acceptable carrier to form a pharmaceutical composition.

The pharmaceutical compositions should be sterile and contain a therapeutically effective amount of the bacteriophages (or target cells containing the bacteriophages) in a unit of weight or volume suitable for administration to a patient. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredients. The term "physiologically acceptable" refers to a non-toxic material that is compatible with a biological system such as a cell, cell culture, tissue, or organism. The characteristics of the carrier will depend on the route of administration. Physiologically and pharmaceutically acceptable carriers include diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials which are well known in the art.

In one particular embodiment, the preferred pharmaceutical composition is contained in an implant that is suitable for implantation into the mammalian recipient. Exemplary bioerodible implants that are useful in accordance with this method are described in PCT International application no. PCT/US/03307 (Publication No. WO 95/24929, entitled "Polymeric Gene Delivery System", claiming priority to U.S. patent application serial no. 213,668, filed March 15, 1994). PCT/US/0307 describes a biocompatible, preferably biodegradable polymeric matrix for containing an exogenous gene under the control of an appropriate promoter. The polymeric matrix is used to achieve sustained release of the exogenous gene in the patient. In accordance with the instant invention, the bacteriophage particles described herein are encapsulated or dispersed within the biocompatible, preferably biodegradable polymeric matrix disclosed in PCT/US/03307. The polymeric matrix preferably is in the form of a micro particle such as a micro sphere (wherein the bacteriophage particle is dispersed throughout a solid polymeric matrix) or a microcapsule (wherein the bacteriophage particle is stored in the core of a polymeric shell). Other forms of the polymeric matrix for containing the bacteriophage particle include films, coatings, gels, implants, and stents. The size and composition of the polymeric matrix device is selected to result in favorable release kinetics in the tissue into which the matrix device is implanted. The size of the polymeric matrix further is selected according to the method of delivery which is to be used, typically injection into a tissue or administration of a suspension by aerosol into the nasal or pulmonary areas. The polymeric matrix composition can be

selected to have both favorable degradation rates and also to be formed of a material which is bioadhesive, to further increase the effectiveness of transfer when the polymeric matrix is administered to a mucosal surface. The matrix composition also can be selected not to degrade, but rather, to release by diffusion over an extended period of time.

Both non-biodegradable and biodegradable polymeric matrices can be used to deliver the bacteriophage particles of the invention to the subject. Biodegradable matrices are preferred. Such polymers may be natural or synthetic polymers. Synthetic polymers are preferred. The polymer is selected based on the period of time over which release is desired, generally in the order of a few hours to a year or longer. Typically, release over a period ranging from between a few hours and three to twelve months is most desirable. The polymer optionally is in the form of a hydrogel that can absorb up to about 90% of its weight in water and further, optionally is cross-linked with multi-valent ions or other polymers.

In general, the bacteriophage particles of the invention are delivered using the bioerodible implant by way of diffusion, or more preferably, by degradation of the polymeric matrix. Exemplary synthetic polymers which can be used to form the biodegradable delivery system include: polyamides, polycarbonates, polyalkylenes, polyalkylene glycols, polyalkylene oxides, polyalkylene terphthalates, polyvinyl alcohols, polyvinyl ethers, polyvinyl esters, poly-vinyl halides, polyvinylpyrrolidone, polyglycolides, polysiloxanes, polyurethanes and co-polymers thereof, alkyl cellulose, hydroxyalkyl celluloses, cellulose ethers, cellulose esters, nitro celluloses, polymers of acrylic and methacrylic esters, methyl cellulose, ethyl cellulose, hydroxypropyl cellulose, hydroxy-propyl methyl cellulose, hydroxybutyl methyl cellulose, cellulose acetate, cellulose propionate, cellulose acetate butyrate, cellulose acetate phthalate, carboxylethyl cellulose, cellulose triacetate, cellulose sulfate sodium salt, poly(methyl methacrylate), poly(ethyl methacrylate), poly(butylmethacrylate), poly(isobutyl methacrylate), poly(hexylmethacrylate), poly(isodecyl methacrylate), poly(lauryl methacrylate), poly(phenyl methacrylate), poly(methyl acrylate), poly(isopropyl acrylate), poly(isobutyl acrylate), poly(octadecyl acrylate), polyethylene, polypropylene, poly(ethylene glycol), poly(ethylene oxide), poly(ethylene terphthalate), poly(vinyl alcohols), polyvinyl acetate, polyvinyl chloride, polystyrene and polyvinylpyrrolidone.

Examples of non-biodegradable polymers include ethylene vinyl acetate, poly(meth)acrylic acid, polyamides, copolymers and mixtures thereof.

Examples of biodegradable polymers include synthetic polymers such as polymers of

lactic acid and glycolic acid, polyanhydrides, poly(ortho)esters, polyurethanes, poly(butic acid), poly(valeric acid), and poly(lactide-cocaprolactone), and natural polymers such as alginate and other polysaccharides including dextran and cellulose, collagen, chemical derivatives thereof (substitutions, additions of chemical groups, for example, alkyl, alkylene, hydroxylations, oxidations, and other modifications routinely made by those skilled in the art), albumin and other hydrophilic proteins, zein and other prolamines and hydrophobic proteins, copolymers and mixtures thereof. In general, these materials degrade either by enzymatic hydrolysis or exposure to water *in vivo,* by surface or bulk erosion.

Bioadhesive polymers of particular interest include bioerodible hydrogels described by H.S. Sawhney, C.P. Pathak and J.A. Hubell in Macromolecules, 1993, 26, 581-587, the teachings of which are incorporated herein, polyhyaluronic acids, casein, gelatin, glutin, polyanhydrides, polyacrylic acid, alginate, chitosan, poly(methyl methacrylates), poly(ethyl methacrylates), poly(butylmethacrylate), poly(isobutyl methacrylate), poly(hexylmethacrylate), poly(isodecyl methacrylate), poly(lauryl methacrylate), poly(phenyl methacrylate), poly(methyl acrylate), poly(isopropyl acrylate), poly(isobutyl acrylate), and poly(octadecyl acrylate). Thus, the invention provides a composition of the above-described bacteriophages for use as a medicament, methods for preparing the medicament and methods for the sustained release of the medicament *in vivo*. In the preferred embodiments, the bacteriophage is selected from the group of bacteriophages shown in Table 1 and is labeled with an appropriate ligand as described above. Preferably, the bacteriophage is selected from the group consisting of a lambda phage, a P1 phage, a P22 phage, and an SPP1 phage; more preferably, the bacteriophage is a lambda phage or a P1 phage; most preferably, the bacteriophage is a lambda phage.

Examples

Introduction to the Examples

In the present method, a bacteriophage particle serves as a vector for gene delivery to a mammalian cell. The bacteriophage capsid and/or tail are modified to contain a ligand (or an adapter, e.g., avidin, for a ligand) that can bind to a receptor contained on the surface of a mammalian cell and facilitate entry therein. Following receptor-ligand interaction the recombinant bacteriophage can be internalized via receptor mediated endocytosis or other receptor-mediated mechanisms of intracellular transport. The bacteriophage genome is engineered to include transcriptional control signals, which allow expression of an exogenous polynucleotide in the mammalian cell. Optionally, the recombinant bacteriophage further contains an endosomal lysis signal, such as that contained within the capsid-stabilizing protein VI and L3/p23 of adenovirus (Greber UF et al. EMBO J, 1996; 15: 1766) to promote export of the bacteriophage genome into the cytoplasm prior to fusion of the endocytosed vesicles with mammalian cell lysosomes and enzymatic degradation of the bacteriophage genome. Conventional recombinant DNA techniques can be used for inserting these and other nucleic acid sequences into the phage genome. For example, the adenovirus endosomal lysis signal can be inserted in-frame into or appended to the bacteriophage tail fiber J gene.

To effect successful transduction after entry of the recombinant bacteriophage into the cytoplasm, the bacteriophage genome is delivered to the nucleus. In dividing cells, the genome of the bacteriophage used in the invention is efficiently delivered to the nucleus during mitosis when the nuclear membrane dissipates and reforms. For delivery of the bacteriophage genome to the nucleus of non-dividing cells, the phage particle is modified to include nuclear localization signals (NLS) that mediate transport of molecules or macromolecular aggregates, such as viral DNA across the nuclear membrane and into the nucleoplasm. Such nuclear localization signals are known to those of ordinary skill in the art and include portions of the human immunodeficiency virus (HIV) proteins matrix (MA) and Vpr (Naldini L et al. Science 1996; 272: 263; Bukrinsky MI et al., Nature 1993; 365: 666), large T antigen of SV40 (Kalderon D et al. Cell 1984; 39: 499; *Drosophila* Antennapedia (Derossi D et al. J Biol Chem 1996; 271: 18188), VP22 of Herpes virus (Elliott G. et al., Cell 1997; 88: 223), X. *laevis* nucleoplasmin, X. *laevis* lamin L, human c-myc encoded protein, and adenovirus type 2/5 Ela (Chelsky D et al. Mol

Cell Biol 1989; 9: 2487). For example, the HIV or SV40 nuclear localization signal can be inserted in frame into or appended to the tail fiber J protein or the H protein which is thought to bind tightly to the bacteriophage genome. Exemplary nuclear localization signals as disclosed in the above-identified references are provided in SEQ. ID NOS. 7-14.

In a preferred embodiment, the bacteriophage genome further includes adeno-associated viral sequences that mediate integration into specific chromosomal regions of the mammalian cell to effect delivery of the transduced bacteriophage genome to a specific mammalian cell chromosomal location. Such adeno-associated viral sequences are known to those of ordinary skill in the art. In yet other embodiments, the bacteriophage has a genome of sufficient size (e.g., P1 phage) to permit the further inclusion of nucleic acid sequences located 5' and 3' of the exogenous polynucleotide to promote site directed homologous recombination. More preferably, the bacteriophage has a genome of sufficient size (e.g., P1 phage) to permit insertion of a large genetic unit containing not only one or more coding sequences of interest and their associated control regions, but in addition, mammalian autonomous replicating sequences, as well as centromere and telomere sequences in a self replicating format to form, in effect, an artificial chromosome. The ability to deliver exogenous genetic material of large size (e.g., a complete gene, including regulatory sequences) has obvious advantages with respect to achieving stable and tissue specific expression *in vivo* and *in vitro*.

A description of several bacteriophage vectors that can be used in accordance with the methods of the invention follows. Briefly, Example I illustrates the construction of three bacteriophage λ vectors containing exogenous genetic material; Example II illustrates the post-packaging modification of bacteriophage λ capsid and tail proteins to form a galactose-labeled bacteriophage λ; Example III illustrates the targeted delivery of a modified bacteriophage of the invention to human hepatoma G2 (HEP G2) cells grown *in vitro*; Example IV illustrates the targeted delivery of a modified bacteriophage of the invention to hepatocytes in the liver of live mice in vivo; Example V illustrates the construction of bacteriophage vectors containing modified tail proteins; and Example VI illustrates screening methods for selecting bacteriophages and target cells that transduce mammalian cells via receptor-mediated endocytosis. These examples are provided for illustrative purposes only, and their inclusion is not meant to limit the scope of invention. Thus, each of the procedures described in reference to the bacteriophage λ can be performed using any of the bacteriophages identified in Table 1 by: (1) obtaining the

published restriction maps for these bacteriophages; (2) selecting an exogenous polynucleotide (e.g., a gene having a published sequence that encodes a therapeutic polynucleotide product, such as identified in Tables 2 or 3); (3) inserting the gene into the well-defined restriction sites of the bacteriophage genome using substantially the same procedures described herein (e.g., blunt-ended ligation into bacteriophage lambda) or alternative procedures known to one skilled in the art for the insertion of a polynucleotide into a restriction site of a bacteriophage; (4) modifying the surface of the bacteriophage to include a ligand that targets delivery of the modified bacteriophage to a mammalian cell that contains on its surface the cognate receptor for the ligand (such as identified in Table 6) using substantially the same procedures described herein (e.g., galactosylation of the bacteriophage lambda phage particles, otherwise chemically modifying the surface of the bacteriophage to attach a ligand, or engineering the bacteriophage to express the ligand as a surface protein or polypeptide); and (5) contacting the modified bacteriophage with the target cell under conditions that permit the selective binding of the bacteriophage ligand to the cognate receptor using substantially the same conditions described herein for *ex vivo* and *in vivo* targeted delivery of bacteriophage lambda or alternative procedures known to one of ordinary skill in the art. By "substantially the same conditions", it is meant that the conditions are modified to substitute a different bacteriophage for bacteriophage lambda in the procedure and to use the published restriction enzymes, buffers, incubation times, size inserts and so forth that are known to be appropriate for modifying the different bacteriophage to include an exogenous polynucleotide and contain a surface ligand.

Example I Construction of three bacteriophage λ vectors

(a) Construction of the CMV-β- gal - lambda DASH II Chimeric DNA Sequences

The CMV β-gal l chimeric DNA sequences used in these experiments was constructed as follows. A 10.7 kilobase pair (kb) Eco RI fragment excised from plasmid adCMV/ NLS β-gal (obtained from Dr. Ronald Crystal, Columbia University; see Figure 3) that contained the reporter gene expression cassette was ligated to Eco RI/ BamHI double-digested left and right arm of λ DASH II (Stratagene, La Jolla, CA), a derivative of λ 2001, which contains an extended range of cloning sites. The genetic and physical map information used in this invention was obtained from published wild type λ DNA sequences and from the GenBank data base (accession number J02459).

The 10.7 kb reporter gene expression cassette included DNA sequences corresponding to

the CMV promoter/ enhancer linked to the E. coli β-gal gene. The expression cassette also included sequences encoding simian virus 40 (SV40) RNA splicing donor / acceptor sites and polyadenylation signals.

Generation of recombinant λ phage particles. CMV β-gal containing recombinant λ phage particles were generated by packaging the ligated CMV β-gal λ DASH II chimeric DNA with an *in vitro* packaging extract, Giga pak Gold II (Stratagene, La Jolla, CA), according to the manufacturer instructions.

Large scale preparation of recombinant phage particles. Conventional methods can be used to propagate the recombinant λ phage (e.g., Meth. Enzymol. 152:145-170, 1987). For example, the bacteriophage λ that was used in this experiment was grown and amplified from a single agar plug. Liquid lysates were made by growing the E. coli strain XL1 blue MR (P2) (genotype Δ (mcr A)) 183Δ (mcr CB- hst MR) 173 end A 1 sup E 44 thi-1 gyr 96 rel A1 lac (P2 lysogen), to A_{550} greater than or equal to 0.5 in Luria broth containing 5mM $CaCl_2$ at 37°C with vigorous shaking. At this point, the phages were added to a "moi" (multiplicity of infection) of 0.01. After 4-5 hrs of vigorous shaking, the lysed culture broth was centrifuged to remove cell debris. Amplified bacteriophages were concentrated by polyethylene glycol (PEG) precipitation from the lysate in accordance with standard procedures. PEG precipitated phages were further purified by $CsCl_2$ gradient centrifugation in accordance with standard procedures. (See, e.g., Meth. Enzymol. 152:145-170, 1987 for standard PEG precipitation and $CsCl_2$ gradient purification protocols.)

Construction of phosphoglycerokinase (PGK) β-gal λ DASH II chimeric DNA sequences. Expression from the CMV promoter/enhancer by mouse liver *in vivo* reportedly is extremely low (Furth, A., et al., Nucleic Acid Research 19:6205-08 (1991)). Accordingly, for *in vivo* targeted delivery of reporter β-gal gene to mouse liver, a chimeric PGK (phosphoglycerokinase) promoter/enhancer/β-gal/ λ DASH II (Fig. 4) was constructed as follows. Eco RI and Bal I digested 0.6 kb fragment released from the PGK promoter containing plasmid (Gene, 80:65, 1987) was subcloned into Eco RI Xho I/ blunt site of adaCMV/ NLS/ beta-gal plasmid. The resulting plasmid (designated PGK-NLS-β-gal) was Eco RI digested and ligated to λ DASH II

sequences. After ligation, the chimeric PGK-NLS-β-gal λ DASH II DNA was packaged *in vitro* and subsequently propagated in E. coli for large scale production of phage particles.

Construction of MCK-DMD-λ DASH II chimeric DNA sequence. As discussed previously, a significant advantage of the bacteriophage λ used in this invention is that it can be readily engineered to permit large gene-containing expression cassettes that could never be achieved by viral vectors currently used for delivery to mammalian cells. An example is a phage vector containing the MCK-DMD gene, which at 20.3 kb (Fig 5) greatly exceeds the capacity of currently used vectors. This vector was constructed as follows. The 20.3 kb expression cassette containing muscle creatine kinase (MCK) promoter/enhancer 5' to the full length DMD cDNA (obtained from J. Chamberlain, U. Michigan) was released from the plasmid pMDA after digestion with BssHII enzyme in accordance with standard procedures. This fragment was blunt ended and ligated to λ DASH II DNA sequences. The MCK-DMD gene / λ DASH II DNA sequences were used for generating the recombinant bacteriophage λ particles after packaging the chimeric DNA sequences with *in vitro* packaging extract (as described above).

Construction of λ phage particles displaying the ligand "L" heregulin for targeted gene delivery Construction of λ phage particle that displays a specific ligand can be used for delivering the gene to a specific cell or tissue expressing the corresponding receptor via receptor/ligand mediated endocytosis. DNA sequences that encode ligand(s) known to be endocytosed after interaction with cognate receptor(s) were inserted, in frame, into the D-gene locus of the λ phage genome to produce a transcription template for a chimeric D gene-ligand bifunctional protein that expresses the fusion protein (including the ligand(s) of choice) on the bacteriophage surface. A similar strategy recently has been reported for targeting retrovirus vectors for gene delivery into a specific cell or tissue type (Proc. Natl. Acad. Sci. 92:9747-51 (1995)). For example, ligand directed retroviral targeting of human breast cancer cells recently has been reported by constructing retrovirus vectors expressing heregulin -gp70 chimeric envelope genes.

Alternatively, ligand directed λ phage targeting of specific mammalian cells is achieved by expressing chimeric λ D-gene and immunoglobulin single chain variable fragments (SCVF) directed against receptors such as LDL and IL-2 , both of which are known to be internalized after the antibody-receptor interaction. Recently this approach has been used to generate

retrovirus vectors expressing chimeric envelope and SCVF for LDL receptor protein to transfer a β-gal gene to human cells expressing the LDL receptor (Proc. Natl. Acad. Sci. 92:7570).

Bacteriophage λ particles having a capsid that displays a chimeric λ D-gene-avidin fusion protein also can be used to effect targeted delivery of the phage particles to a specific mammalian cell. Since the avidin molecule has very high affinity for biotin, any peptide or polypeptide ligand that can be biotinylated without adversely affecting the ability of the receptor to mediate endocytosis can be used as described herein. The avidin gene, or portion thereof encoding the polypeptide that selectively binds to biotin, is inserted, in frame, into the bacteriophage genome using the procedures described herein for forming a chimeric λ D-gene and the published cDNA nucleic acid sequence for avidin (Gope, L. Mohan, et al., Nucleic Acid Research 15:3595-3606 (1987) and GenBank Accession No. 451889).

Fig. 6 illustrates the construction of a λ phage particle expressing the heregulin-D-gene chimeric DNA sequences on its capsid. In this construction, the D gene heregulin sequences were first fused by generating 3 fragments (AB, CD, and EF) by polymerase chain reaction (PCR) in accordance with standard procedures using Taq polymerase and 6 sets of primers (a,b,c,d,e,f):

primer a (SEQ. ID NO. 1) ATACCGAGGGCTGCAGTGTACA

primer b (SEQ. ID NO. 2) CTCTTTCAATTGGGGAGGCAAAACGATGCTGATTGCCGTTC

primer c (SEQ. ID NO. 3) TTGCCTCCCCAATTGAAAGAG

primer d (SEQ. ID NO. 4)
GTGATGAAGGGTAAAGTTATTTGCGTTTTTTTTTCGGCGGGGTCCTCCATAAATT
CAATC

primer e (SEQ. ID NO. 5)
TAACTTTACCCTTCATCACTAAAGGCC

primer f (SEQ. ID NO. 6)
AAACGTACAGCGCCATGTTTACCAG

Gel purified fragments AB, CD and EF were treated with T4 polymerase to remove the overhanging nucleotide "A" in the fragment generated by terminal transferase activity present in the Taq polymerase used for PCR amplification. This step ensures the joining "in frame" of the coding sequences of both the λ D-gene and heregulin. Next,

fragments AB and CD were joined by first denaturing and then annealing the partial overlapping sequences, followed by extension with Taq polymerase, and then amplification after the addition of primers a and d. The fragment AD was gel purified, treated with T4 polymerase, and subsequently used for joining fragment EF (as described above). The fused λ D-gene/heregulin gene containing fragment AF was subcloned into the TA cloning vector (Invitrogen, San Diego, CA). The clones containing the AF fragment were identified by determining the sequences with a double-stranded sequencing method using the Sequenase 2.0 kit (USB, Cleveland, OH). The AF fragment was generated after digestion of the TA plasmid clone with the restriction enzyme BsrGI and ligated into the necessary fragments of λ (as shown in Fig. 6). This chimeric DNA construct was used to generate λ phage particles having a genome that includes the recombinant D-gene/heregulin construct and a capsid that displays this chimeric protein.

Example II

Post-packaging modification of bacteriophage λ capsid and tail proteins to form a galactose-labeled bacteriophage λ

Chemical Modification.

The particular procedure for modifying the bacteriophage λ capsid and tail proteins to form a galactose-labeled bacteriophage λ is described herein. The galactose-labeled bacteriophage particles can be used to target liver cells which express a unique asialo-glycoprotein receptor. This procedure is based upon the published procedures for forming an artificial asialo-glycoprotein containing lactose (Neda, H., et al., JBC 296:14143-14146 (1991)) or galactose (Human Gene Therapy 5:429-435 (1994)).

Method. $CsCl_2$ purified CMV β-gal reporter gene containing λ phage particles (10^{11} PFU/ml) were galactosylated in 2 ml reaction volume containing 60mg of galactose and 100 mg of 1 ethyl-3-(3-diethylaminopropyl) carbodiaminide (EDC, Sigma Chemical. Corp., Saint Louis, MO) in sodium chloride solution. The pH of the unbuffered solution was adjusted to 7.5 with NaOH and the reaction mixture was incubated at room temperature for various lengths of time ranging from 24 to 48hrs. The galactose associated with phage particles was determined using ^{14}[C]-labeled galactose. Twenty five μl of ^{14}[C]-galactose (250-360mCi/mM, DuPont NEN, Boston, MA) were included

in a 100:1 reaction mixture except the cold galactose was omitted. At the end of the different incubation periods, the solutions were filtered and washed through nitrocellulose filters in a vacuum filtration device. The filters were removed and counted in a Beckman scintillation counter (Palo Alto, CA). The radioactivity on filters was converted to numbers of ^{14}[C]-galactose on the basis of the specific activity. The extent of galactosylation was expressed as numbers of galactose/phage particles.

As described in the detailed description of the invention, a variety of chemical methods can be used to attach a ligand to bacteriophage capsid and/or tail proteins. For example, small peptide ligands, such as insulin, epidermal growth factor (EGF), keratinocyte growth factor (KGF), Fab fragments for anti-polymeric immunoglobulin receptors can be covalently linked to λ phage using the hetero-bifunctional crosslinking reagent N-succinimidyl 3-(2-pyridyl dithio) propionate (SPDP) or other bifunctional crosslinking agents in accordance with manufacturer's instructions. Additional crosslinking agents are provided in the description and are known to those of ordinary skill in the art. After coupling the ligand to the bacteriophage surface proteins, the modified bacteriophage is used to selectively deliver the gene of interest to target cells that express a cognate receptor for the ligand (Example III).

Example III

Targeted delivery of a Modified Bacteriophage of the Invention to human hepatoma G2 (HEP G2) cells grown *in vitro*

In vitro targeted delivery of CMV β-gal expression cassette containing galactose modified λ phage particles.

1. Growth of cells:

Conventional tissue culture methods were used to grow HEP G2 cells. HEP G2 cells were cultured in minimal essential medium as modified by Eagle (EMEM) containing 10% FBS. Cells were seeded one day prior to the addition of λ phage particles for gene transfer experiments. *In vitro* targeted delivery of the β-gal gene to HEP G2 cells was accomplished by allowing the phage particles to interact with the cells in tissue culture growth medium for about 6-10 hrs; more preferably, for 8-10hrs. In general, after galactosylation, 10^7-10^{10} total phage particles/ml for 6-10 hrs, preferably 10^7 particles/ml for 10 hours, are needed for efficient transduction of targeted cells. After

exposing the cells to phage particles for the appropriate time, the phage-containing medium was removed and replaced with fresh media.

2. Detection of internalization of bacteriophage vectors and gene expression:

After ligand-receptor mediated endocytosis of a bacteriophage vector into a mammalian cell, the expression of the exogenous genetic material in the mammalian cell can be monitored using standard methodologies. For example, delivery of a bacteriophage modified with a CMV β-gal recombinant construct and a galactose ligand to a HEP G2 cell receptor can be measured by detecting bacteriophage DNA or RNA by Southern or northern blotting or *in situ* hybridization with or without amplification by polymerase chain reaction (PCR). Radioactively-labeled DNA or RNA probes that selectively hybridize to unique portions of the phage genome, e.g., portions corresponding to a regulatory sequence, such as a promoter, or portions corresponding to the exogenous genetic material, such as the reporter gene β-galactosidase, can be designed and constructed using standard molecular biology techniques. Where the modified bacteriophage is to be used for delivering exogenous genetic material to a mammalian cell *in vivo* (e.g., to hepatocytes in murine liver), delivery of the phage to the specific cell can be detected by obtaining the targeted cells in a biopsy and assaying the biopsied cells using, for example, the above- mentioned conventional methods (e.g., Southern or northern blotting).

In general, expression of a gene of interest (e.g., therapeutic or marker polynucleotide) in a mammalian cell is detected by measuring the functional or immunological activity of the expressed gene in the targeted tissue, targeted cell, or body fluid (e.g., serum, lymph fluid). Alternatively, direct RNA or protein analysis for detecting specific transcription or translation products can be performed in accordance with standard practice. Exemplary detection techniques to measure gene expression include one or more of the following techniques, alone or in combination: northern or western blotting, *in situ* hybridization, reverse transcription, PCR amplification, immunostaining, RIA and ELISA. Such routine techniques also can be used to measure the stability of expression and maintenance of the delivered exogenous genetic material in the mammalian target cell by, for example, measuring the expression of one or more

reporter (e.g., marker) genes as a function of time following transduction. In an exemplary protocol, a reporter gene #1 initially is delivered to the target mammalian cells. One week later, reporter gene #1 expression is assayed. One to three months following the initial transduction, a reporter gene #2 is transduced into the same population of target cells. One week later, the ratio of reporter gene #1 to reporter gene #2 is determined by, for example, immunostaining (e.g., using a different dye for each immunohistologic assay) the target tissue to assess the stability and maintenance of expression of the exogenous genetic material in the target cell population. A diminution of reporter gene #1 expression compared to reporter gene # 2 expression can indicate either immune attack against transduced cells or transduction of short-lived, differentiated cells. An exemplary procedure for processing a target tissue to which a β-galactosidase reporter gene expression has been delivered to determine whether the reported gene is expressed therein includes: a) snap freezing the tissue in isopentane chilled with liquid nitrogen, b) mounting the tissue on cryomold (Tissue -Teck, Miles, Elkhart, IN) using OCT and freezing, c) cutting the frozen tissue with a microtome at -20°C into 10um sections , d) staining the tissue for β-galactosidase reporter gene expression with X-gal (1mg/ml) (a reporter gene substrate, Boehringer Mannheim, Indianapolis, IN), potassium ferro- and ferricyanide (35mM each) in phosphate buffered saline solution. To detect expression of a therapeutic polynucleotide, the above-described procedure is used with the exception that immunostaining using an antibody to detect the therapeutic polynucleotide expression product is used in place of the reporter gene substrate assay.

In the following example, delivery of the CMV β-gal gene-containing galactosylated and ungalactosylated phage particles *in vitro* to HepG2 cells and *in vivo* (Example IV) to liver cells of mice was measured. Histochemical staining of HepG2 cells using X-gal was used to measure expression of the β-gal reporter gene. The results showed a field of Hep G2 cells expressing the β-gal gene, as indicated by the appearance of positive cells that stained darkly following incubation with the substrate X-gal. The control experiment (ungalactosylated phage particles) did not show positive cells that stained darkly following incubation with the substrate X-gal. These results demonstrate that the galactose-labeled phage particles were selectively internalized by the murine liver cells, presumably via receptor-mediated endocytosis.

Example IV

The targeted delivery of a modified bacteriophage of the invention to hepatocytes in the liver of live mice *in vivo*.

For the in vivo experiment, 10^{10} phage particles, either modified with an average of 50 galactose residues/phage particle or unmodified (control), were injected into mice intraperitoneally. Three days after injection, mice were sacrificed and liver and kidney tissue biopsies were prepared. Frozen sections were cut and fixed with 1% glutaraldehyde and subsequently stained with X-gal. After staining, tissue sections were analyzed by light microscopy. Unmodified phage (control) injection did not result in β-gal expression in either liver or kidney tissue sections that were stained with X-gal as detailed above. In contrast, injection of galactosylated phage resulted in detectable β-gal staining in hepatocytes but not in any cells on kidney tissue sections from the same mouse. Careful analysis of 40 serial sections of liver removed from the mouse exposed to the galactosylated phage demonstrated that approximately 0.5 to 1% of hepatocytes stained with X-gal and thus, effectively were transduced by the recombinant bacteriophage vector. While this example shows that intraperitoneal administration could be used successfully to transduce hepatocytes *in vivo*, one skilled in the art reasonably would believe that alternative routes of administration (such as those described in the detailed description), as well as the further inclusion of endosomal lysis signals and/or nuclear localization signals also can be used to successfully transduce mammalian target cells *in vivo* and *in vitro*.

Example V

Construction of Bacteriophage Lambda vectors with Modified Tail Proteins.

As described above, bacteriophage with wild type packaging proteins and the beta-galactosidase gene are not effective genetic vectors for transducing Hep G2 cells. However, chemical modification of the surface proteins in vitro with galactose such that it will bind to and be internalized by the galactose-receptor on the Hep G2 cell surface does result in an effective vector. The galactose-asialoglycoprotein ligand-receptor interaction and subsequent internalization via receptor mediated endocytosis have been described in detail in the literature. Accordingly, the above-noted results, together with the literature reports documenting the mechanism of asialoglycoprotein-mediated transport, supports our hypothesis that the transfer of genetic material from the above-described modified bacteriophages to a mammalian cell target occurs via receptor-mediated endocytosis.

Thus, we believe that the mechanism by which modified bacteriophage vectors transfer genetic material to mammalian target cells is fundamentally different from the mechanism by which wild type bacteriophages transduce their natural prokaryotic hosts and that functional bacteriophage tail proteins are not essential for mammalian target cell transduction using the bacteriophages of the invention. This hypothesis is confirmed using the screening methods described below.

Using standard methods, endocytosis inhibition agents (e.g., colchicine, cytochalasin B and D, and monodansylcadaverine) are used to arrest cell-mediated endocytosis in a mammalian cell and, thereby, select for bacteriophages whose mechanism of transfer does not depend on injection of genetic material into the mammalian call as is observed in the naturally occurring bacteriophage transduction of prokaryotic hosts. Alternatively, antibodies that bind to bacteriophage tail proteins and block their function can be used to demonstrate that wild type tail function is not required for the transfer of the bacteriophage genome to the host cell nucleus. The methods also are useful for identifying target cells that contain an appropriate receptor in sufficient quantity to internalize the preferred bacteriophage vectors via receptor-mediated endocytosis. The screening assays disclosed herein demonstrate that bacteriophage vectors which include one or more non-functional tail protein(s) and a standard genomic expression marker, e.g. beta-galactosidase or other detectable protein, are incapable of transferring genetic material to the target cells in the presence of the endocytosis inhibition agents but are capable of transduction when contacted with the target cells in the absence of such inhibition agents. Thus, the assay is useful for identifying novel bacteriophages which are structurally and functionally distinct from naturally-occurring bacteriophages and modified bacteriophage such as those described in PCT publication no. WO 96/21007, entitled, "Bacteriophage-mediated Gene Transfer Systems Capable of Transfecting Eukaryotic Cells", applicant Chiron Viagene, Inc. which report modified bacteriophages that inject their genetic contents into mammalian target cells.

Standard recombinant methods and screening technology are used to prepare a preferred bacteriophage vector with genetically modified tail protein genes that either contain amber mutations or sequences that enhance binding of the bacteriophage vector to the mammalian cell surface and subsequent internalization via receptor-mediated

endocytosis. Exemplary essential tail proteins in the lambda phage that can be modified to prepare preferred embodiments of the invention include: H, J, M proteins. As used herein, "essential tail proteins" refers to those proteins that are essential for facilitating the injection of the bacteriophage genetic material into its natural prokaryotic host. Although Applicants doubt the likelihood that bacteriophage can be selected or designed to have tail proteins that are capable of facilitating the injection of the bacteriophage genetic material into a eukaryotic host, the phrase "essential tail proteins" is also meant to embrace the tail proteins of such hypothetical eukaryotic cell-injecting bacteriophage. One or more of these or other tail proteins can be rendered non-functional (i.e., incapable of facilitating the injection of the genetic material into the host cell) using recombinant, mutagenesis, and/or chemical methods in accordance with procedures known to one of ordinary skill in the art. Such procedures can, of course, be applied to the preparation of other types of modified bacteriophages by, e.g., identifying the essential tail proteins of one or more bacteriophages in Table 1 and modifying the essential tail proteins as described herein.

The modifications to the essential tail protein(s) renders the bacteriophage incapable of mediating the transfer of genetic material into a mammalian host via an injection mechanism analogous to that responsible for the naturally occurring bacteriophage transduction of prokaryotic hosts. Preferably, such modification(s) of the tail proteins further render the bacteriophage incapable of infecting and/or propagating within its natural prokaryotic host (Su^0 E.Coli bacteria). Accordingly, such tail mutant bacteriophage vectors are packaged in vitro using specific purified protein packaging extracts and recombinant bacteriophage genomes or in Su^- E. Coli bacteria (for amber mutant containing bacteriophage vectors). The benefit of such preferred bacteriophage vectors is that these vectors are incapable of propagating in the host organism's natural flora.

In certain embodiments, the tail proteins are modified so that they function to properly package the virus' genome but have lost their ability to mediate injection across bacterial cell membranes. Phages containing these modified tail-proteins require packaging in a cell-free system as described above.

In yet another embodiment, the mutations in the tail proteins are temperature sensitive such that at temperatures other than mammalian physiologic temperature, the tail

protein functions normally and can mediate prokaryotic cell infection; however, at mammalian physiologic temperature, the temperature sensitive mutation inactivates the wild type tail function. Such mutant bacteriophage vectors can be propagated in a prokaryotic host and packaged using standard procedures.

In a further embodiment, the tail protein is modified to include signals that target and/or facilitate entry of the vector's genetic material into subcellular organelles, including the nucleus. Exemplary signals for targeting the nucleus include polypeptides derived from the matrix or Vpr proteins of HIV or the large T-antigen of SV40. Such signals can be inserted into the lambda or other bacteriophage genomes using standard procedures. Preferably, such signals are integrated into the lambda phage genome in frame either within or appended to the J or H genes. For example, the genome of the modified bacteriophage described in Example 3 could be further modified in the following way. First using PCR fusion, a DNA fragment containing the SV40 nuclear localization sequence coding for NH_2-Pro-Lys-Lys-Lys-Arg-Lys- Val (PKKKRKV)(Kalderon D et al. Cell 1984;39: 499). (SEQ. ID. No. 7) flanked both 5' and 3' by wild type lambda DASH II J gene sequences including Bst 1170 I restriction sites is generated. Then this fusion DNA product is cut by Bst 1170 I leaving the following Bst 1170 I sticky end-5' J gene sequence-SV40/NLS-3' J gene sequence-Bst 1107 I sticky end. This digestion product is ligated, in frame, into the wild type Bst 1107 I restriction site within the J gene at base pair number 18834 by standard recombinant technology. Such signals can be inserted anywhere in the bacteriophage packaging protein coding sequence provided that such insertion does not interfere with receptor-mediated endocytosis and/or expression of the exogenous polynucleotide within the target cell. Optionally, insertion of these signals into the bacteriophage also inactivates wild type function and, optionally, further serves to enhance bacteriophage penetration of the outer membrane of the mammalian cell.

Example VI

Screening Method for Selecting Target Cells and Modified Bacteriophages of the Invention that Transduce Mammalian Target Cells via Receptor-mediated Endocytosis.

The following assay is useful for selecting target cells and bacteriophage vectors that depend on receptor mediated endocytosis for transfer of the bacteriophage genome to the target cell nucleus. The modified bacteriophages are grown to high titer 10^{12} to 10^{13} by standard methods employing either endogenous packaging in a prokaryotic host or in vitro packaging with proteins extracted from appropriately engineered lysogen strains. Additionally or alternatively, these modified bacteriophages may be modified chemically after packaging using standard methods. The coding sequence for beta-galactosidase or other easily detectable gene expression indicator(s) is incorporated into the genome of the test bacteriophages using standard procedures. At preferable multiplicities of infection of 1 to 100, purified bacteriophage is then added to sub-confluent cultures on cover slips of mammalian test cells that bear the receptor to which a modified bacteriophage has been targeted. One example of such a test cell line is HepG2 cells in which the asialoglycoprotein mediated endocytosis of galactose bearing oligosaccharides and macromolecules is well studied (described above).

The screening assay is performed in the presence and absence of endocytosis inhibitors. Preferably, prior to addition of the modified bacteriophage, inhibitors of endocytosis are added to the cell medium containing the target cells either singly or in various combinations at concentrations such that receptor mediated endocytosis is optimally inhibited without irreversibly injuring the test target cells. Exemplary endocytosis inhibitors include colchicine, taxol, monodansylcadaverine, cytochalasin B, or cytochalasin D. As a control, cells not treated with endocytosis inhibitors are infected with test bacteriophages in parallel.

The test cells are incubated with bacteriophage for about 48 hours or other suitable time to allow transduction and expression. At various time points during this period, the plated cells are fixed for x-gal staining, immunostaining, in situ hybridization, electron microscopy or other standard methods that are known to one skilled in the art for detecting the inserted exogenous polynucleotide or its expression product. X-gal staining is used to indicate whether or not the transduced genetic material (including a beta-galactosidase marker) is expressed in the test cells. Alternatively, immunostaining with specific antibodies or anti-sense in situ hybridization probes employed to detect either the translation or transcription of a gene inserted into the vector's genome. Additionally or

alternatively, antibodies with specific affinity for the modified bacteriophage surface protein(s) are used to detect the localization of viral coat proteins within test cells, including organelles within the test cells, using confocal immunofluorescence microscopy. Optionally, transmission electron microscopy is employed to determine the sub cellular localization of modified bacteriophages at various time points after introduction of the modified bacteriophage into the test cell media.

It is anticipated that less than about 5% of target cells that are contacted with the endocytosis inhibitor(s) demonstrate internalization and/or expression of the indicator gene. In addition, immunostaining with anti-coat antibodies of sections of test cells that have been exposed to endocytosis inhibitors and fixed directly following incubation with bacteriophage detect the presence of viral coat proteins in a sub-plasma membrane location, thus, indicating entrapment of these bacteriophages within endocytic vesicles prevented from fusion with lysosomes and normal trafficking. In contrast, the bacteriophage-infected test cells in which receptor-mediated endocytosis is not inhibited exhibit expression of the indicator gene in greater than about 15% of the test cells. Furthermore, immunostaining uninhibited test cell sections for coat proteins demonstrate a predominance of sub plasma membrane localization. These results are surprising and unexpected in view of the internalization mechanism proposed in PCT publication no. WO 96/21007 for bacteriophage transfer of genetic material to a mammalian cell.

Example VII:
Construction of a modified bacteriophage vector which displays cyclic RGD ligand on its surface

We have designed a bacteriophage genome which appends, in frame, the coding sequence cyclic RGD ligand 3' to the coding sequence of the native D-gene. This bacteriophage genome was packaged *in vitro* and propagated in *E. Coli* so as to produce productive phage which contains the D-gene-RGD fusion protein product on its surface. The effect of displaying this cyclic RGD ligand on the surface is to promote uptake of this bacteriophage by mammalian cells bearing the alpha-5/beta-1 integrin receptor via cell-mediated endocytosis. This modified bacteriophage is referred to as lambda DASH II-RGD. Using a beta-galactosidase reporting construct, we have found the lambda DASH II-RGD to infect 30-40% of cultured monolayer bovine endothelial cells at a multiplicity

of infection of 1000 phage particles per cell. The sequences of the RGD primers are shown in SEQ ID NOS. 18, 19, 20, and 21.

Example VIII:

Delivery and expression of human factor VIII/von Willebrand factor containing vector

As further enabling proof of our technology, we have constructed a bacteriophage genome which contains the following functional elements in 5' to 3' tandem array: CMV promoter-human factor VIII cDNA (Seq ID No. 15) - internal ribosomal entry site (IRES)-human von Willebrand factor cDNA (Seq. ID No. 17). The IRES cassette was obtained from Clontech (Palo Alto, CA). The total size of this linear DNA construct is 20.6 kilobases. As described in the detailed description of the invention, our FVIII/vWF construct has been spliced into the multiple cloning site of the bacteriophage lambda DASH II-RGD genome via blunt ended ligation (see included sequences and construct diagram).

This resulting genome has been efficiently packaged and propagated to titers of 10^{11} to 10^{12} phage particles per 2 liter broth culture. We have infected subconfluent monolayer cultures of bovine endothelial cells with this FVIII-vWF containing vector. Using a standard assay for factor VIII activity (COATEST VIII: C4-Chromogenenix Mölndal, Sweden), we have measured expression of factor VIII at 48 hours after infection. A 24 hour secretion study revealed that these transduced endothelial cells produce 56 mU/24 hours/10 cm culture plate (approximately $5x10^5$ cells). This secretion rate is 8 fold greater than that reported for a transfected COS cell line (Toole et al., Nature, 312, 1984, p. 342-7).

Example IX: Delivery and expression of a murine dystrophin containing vector

As described in the examples of the patent application, we have obtained a 20.5 kilobase insert from Dr. Jeffrey Chamberlain (Michigan University) containing the following functional elements in tandem oriented 5' to 3': muscle creatine kinase promoter - murine full length dystrophin cDNA (Seq ID No. 17). We, in turn, have spliced this dystrophin insert into lambda DASH II-RGD via blunt end ligation into the multiple cloning site. The resulting recombinant bacteriophage genome has been packaged and

propagated efficiently to titers of 10^{11} to 10^{12} phage particles per 2 liter broth culture.

To test the expression capacity of this dystrophin bacteriophage vector, we injected 50 microliters of 10^{10} phage particles/mL into the gastrocnemius muscle of a dystrophin-deficient mouse (*mdx*). The phage vector was co-injected with a sublethal dose of India ink. This permitted accurate localization of the myocytes in proximity to the needle tract. Employing a polyclonal antibody against dystrophin (Cox et al. Nature, 264, 1993, 725-9), we have been able to demonstrate by immunofluorescence restoration of perimyocyte dystrophin expression. As previously reported (Ibid,), immunofluorescence study of the control uninjected *mdx* mouse muscle with the same anti-dystrophin antibody did not demonstrate specific perimyocyte immunoreactivity.

Each of the references, patents and patent publications disclosed in this document is incorporated in its entirety herein by reference.

While the invention has been described with respect to certain embodiments, it should be appreciated that many modifications and changes may be made by those of ordinary skill in the art without departing from the spirit of the invention. It is intended that such modification, changes and equivalents fall within the scope of the following claims.

The Tables are presented below and are followed by the Sequence Listing and what is claimed:

Table 1 **Exemplary Bacteriophages**

Lambda phage, p1 phage, T even and T odd phages (e.g., T1, T2, T3, T4, T5, T6 and T7); P2; P4; Mu; PM2; N4; SPO1; PBS1; PBS2; 29; SPP1; 6; PR4; PRD1; AP50; DS6A; I3; NS11; Dp-1; MVL2; CP-1; 434; cbk; G; D108; and P7.

Table 2

Exemplary Human Gene Therapy Protocols Approved by RAC

Disease	Gene Therapy Treatment	RAC OK
Severe combined immune deficiency (SCID) due to adenosine deaminase (ADA) deficiency	Autologous lymphocytes transduced with human ADA gene	7/31/90

Advanced cancer	Tumor-infiltrating lymphocytes transduced with tumor necrosis factor gene	7/31/90
Advanced cancer	Immunization with autologous cancer cells transduced with tumor necrosis factor gene	10/07/91
Advanced cancer	immunization with autologous cancer cells transduced with interleukin-2 gene	10/07/91
Familial hypercholesterolemia	*Ex vivo* gene therapy	10/08/91
Malignancy	*In vivo gene* transfer into tumors	2/10/92
Cancer	Gene transfer	2/10/92
Relapsed/refractory neuroblastoma	Cytokine-gene modified autologous neuroblastoma cells (Phase I study)	6/01/92
Brain tumors	Intratumoral transduction with thymidine kinase gene and intravenous ganciclovir	6/01/92
Metastatic melanoma	Immunization with HLA-A2 matched allogeneic melanoma cells that secrete interleukin-2	6/02/92
Advanced renal cell carcinoma	Immunization with interleukin-2 secreting allogeneic HLA-A2 matched renal-cell carcinoma cells	6/02/92
Cancer	Interleukin-4-gene modified antitumor vaccine (pilot study)	9/15/92
Cystic fibrosis	Replication deficient recombinant adenovirus carrying cDNA of normal human cystic fibrosis transmembrane conductance regulator (CFRT) gene; single administration to the lung (Phase I study)	12/03/92
Cystic fibrosis	El-deleted adenovirus vector for delivering CFTR gene (Phase I study)	12/03/92

Disease	Gene Therapy Treatment	RAC OK
Cystic fibrosis	Adenovirus vector used for delivering CFTR gene to nasal epithelium	12/04/92
Recurrent glioblastoma (brain tumor)	*In vivo* tumor transduction using herpes simplex thymidine kinase gene/ganciclovir system	3/01/93
Metastatic renal cell carcinoma	Injection of non-replicating autologous tumor cells prepared +/- granulocyte-macrophage colony stimulating factor transduction (Phase I study)	3/01/93
Cystic fibrosis	Use of replication deficient recombinant adenovirus vector to deliver human CFTR cDNA to the lungs (Phase I study)	3/02/93
Cystic fibrosis	Use of El-deleted adenovirus for delivery of CFTR gene to nasal cavity (Phase I study)	3/02/93
Disseminated malignant melanoma	Human gamma-interferon transduced autologous tumor cells (Phase I study)	6/07/93
Ovarian cancer	Use of modified retro viruses to introduce chemotherapy resistance sequences into normal hematopoietic cells for chemoprotection (pilot study)	6/07/93
Cancer	Immunotherapy by direct gene transfer into tumors	6/07/93
Gaucher's disease	*Ex vivo* gene transfer and autologous transplantation of CD34 + cells	6/07/93
Gaucher's disease	Retro viral-mediated transfer of cDNA for human glucocerebrosidase into hematopoietic stem cells	6/07/93
Asymptomatic patients infected with HIV-1	Murine Retro viral vector encoding HIV-1 genes [HIV-IT(V)]	6/07/93

AIDS	Effects of a transdominant form of *rev* gene on AIDS intervention	6/07/93
Recurrent pediatric malignant astrocytomas	*In vivo* tumor transduction with herpes simplex thymidine kinase gene	6/08/93
Advanced cancer	Human multiple-drug resistance (MDR) gene transfer	6/08/93
Brain tumors	Episome-based antisense cDNA transcription of insulin-like growth factor I	6/08/93
Small-cell lung cancer	Cancer cells transfected with and expressing interleukin-2 gene (Phase I study)	9/09/93

Disease	Gene Therapy Treatment	RAC OK
Breast cancer (post-chemotherapy)	Retro viral mediated transfer of the human MDR gene into hematopoietic stem cells (autologous transplantation)	9/09/93
Recurrent pediatric brain tumors	Intra-tumoral transduction with thymidine kinase gene and intravenous administration of ganciclovir	9/09/93
Malignant melanoma	Immunization with interleukin-2 secreting allogeneic human melanoma cells	9/10/93
HIV infection	Autologous lymphocytes transduced with catalytic ribozyme that cleaves HIV-1 RNA (Phase I study)	9/10/93
Metastatic melanoma	Genetically engineered autologous tumor vaccines producing interleukin-2	9/10/93
Leptomeningeal carcinomatosis	Intrathecal gene therapy	12/02/93
Colon carcinoma	Injection with autologous irradiated tumor cells and fibroblasts genetically modified to secrete interleukin-2	12/2/93
Gaucher's disease	Retro virus-mediated transfer of cDNA for human glucocerebrosidase into peripheral blood repopulating patients' cells	12/3/93
HIV infection	Murine Retro viral vector encoding HIV-IT(V) genes (open label Phase I/II trial)	12/03/93
Advanced (stage IV) melanoma	Induction of cell-mediated immunity against tumor-associated antigens by B7-transfected lethally irradiated allogeneic melanoma cell lines (Phase I study)	12/03/93

Advanced colorectal carcinoma	Immunotherapy by direct gene transfer into hepatic metastases (Phase I study)	12/03/93
Melanoma	Adoptive immunotherapy with activated lymph node cells primed *in vivo* with autologous tumor cells transduced with interleukin-4 gene	12/03/93
Cystic fibrosis	Cationic liposome-mediated transfer of CFTR gene into nasal airway (Phase I study)	12/03/93
Cystic fibrosis	Adenovirus-mediated transfer of CFTR gene to the nasal epithelium and maxillary sinus	12/03/93
Pediatric neuroblastoma	Immunization with gamma-interferon transduced neuro blastoma cells (*ex vivo*) (Phase I)	3/03/94

Table 3

Preferred Therapeutic Polynucleotides and Corresponding Medical Conditions

Dystrophin-Duchenne muscular dystrophy
Globin gene complex-Hemoglobinopathies, e.g. sickle cell anemia, thalassemias
Clotting factor VIII-Hemophilia A
von Willebrand's factor-von Willebrand's disease
Collagen type VII-Epidermolysis bullosa dystrophica
Combinations of cyclin dependent kinase inhibitors, and other cell cycle inhibitors, e.g. p15, p16, p18-Neoplastic processes, e.g. melanoma
Fibrillin-Marfan's syndrome
Polypeptide antigens-Vaccines for tumors, infectious agents
Combinations of cytokines or co-stimulatory immune modulators, e.g., IL-1, IL-2, IL12, GM-CSF, TNFα, IL4, B7-Neoplastic processes
Thymidine kinase-Suicide gene for neoplastic, hyperplastic or hypertrophic processes
Combinations of ribozymes-Targeted against disease predisposing MHC genes or against disease associated messenger RNAs of viral origin, e.g. E6, E7 oncoproteins in HPV, reverse transcriptase in HIV
Individual or combinations of chemotherapy resistance genes to protect bone marrow stem cells from chemotherapy regimens
Polypeptide antigens in conjunction with tolerance inducing sequences, e.g. ribozyme against B7-1-Treatment for autoimmune disease, e.g. rheumatoid arthritis, psoriasis, multiple sclerosis, alopecia areata
Combinations of ribozyme(s), antisense RNA(s), or polypeptide coding sequences(s) for biopolymers that interfere with human viral infections, e.g. HIV, CMV, Hepatitis B and C, Herpes Simplex I, II, HHV-8, EBV, HTLV-I
Combinations of ribozyme(s), antisense RNA(s), or polypeptide coding sequences(s) for biopolymers that interfere with chronic bacterial or parasitic infections, e.g. leprosy, tuberculosis, antibiotic resistant bacteria (e.g., MRSA, VRE), trypanosomiasis, filaraisis, and the like

Table 4

Exemplary Target Cells

Hepatocytes
Melanocytes
Keratinocytes
Myocytes
Adipocytes
Hematopoietic cells, e.g. lymphocytes, erythrocytes, leukocytes, monocytes, progenitor and stem cells
Neurons
Glial cells
Antigen presenting cells, e.g. macrophages, B-cells, Langerhan's cells
Chondrocytes
Osteocytes
Osteoclasts
Endothelial cells
Phagocytes
Fibroblasts
Smooth muscle cells
Renal tubule cells
Mesangial cells
Thymocytes
Bronchopulmonary, gastrointestinal, breast, genitourinary, corneal, renal ductular epithelial cells
Endocrine, exocrine gland cells
Plasma cells
Mast cells
Lens epithelial cells
Retinal epithelial cells
Malignant cells of any derivation
Placental cells
Gonadal cells
Embryonic cells
Zygotes

Table 5

Exemplary Promoters and Enhancers

Constitutive
Phosphoglycerokinase
Long terminal repeat (LTR) of retroviruses, e.g. Moloney murine leukemia virus,
Rous sarcoma virus
Cytomegalovirus promoter

Hematopoietic cells
Promoters
c fms (monocytes, trophoblasts)
T-cell receptor
Enhancers
Immunoglobulin heavy chain
Locus control region of the globin gene complex CD2

Hepatocytes
Promoters
Albumin
α-1-antitrypsin
Pyruvate kinase
Phosphoenol pyruvate carboxykinase
Transferrin
Transthyrethrin
α-fetoprotein
α-fibrinogen
β-fibrinogen
Hepatitis B
Enhancers
Hepatitis B
Tyrosine aminotransferase

Cardiac myocytes
Promoter
Myosin light chain-2
β-myosin heavy chain (cardiac and slow twitch skeletal)
α-cardiac myosin heavy chain
Cardiac alpha actin
Enhancer
β-myosin heavy chain

Fibroblasts
Promoter
Collagen alpha-2 (I)
Elastin (fibroblasts and smooth muscle cells)

Neurons
Promoter
Peripheral myelin protein-22

Adipocytes
Promoter
Lipoprotein lipase
Aromatase cytochrome P450 (adipocytes, brain, ovary)

Thyroid
Promoter
Thyroglobulin

Lens epithelium
Promoter
Crystallin

Breast epithelium
Promoter
Milk protein gene

Skeletal muscle
Promoter
Glut-4
Muscle creatine kinase (skeletal and cardiac muscle)
Enhancer
Muscle creatine kinase (skeletal and cardiac muscle)

Urinary bladder
Promoter
UroplakinII

Keratinocyte
Promoter
Keratin 14
Keratin 10
Involucrin

Melanocyte
Promoter
Tyrosinase

Non specific enhancer elements

SV40
CMV
LTR

Inducible or repressible promoter systems

Estrogen-Gal4 inducible system
RU486-Gal4 inducible system
Tetracycline inducible system
IPTG system
Metallothionein
Tetracycline repressible system

Table 6

Exemplary Receptors and Preferred Ligands

Hepatic receptors
hyaluronic acid
collagen
N-terminal propeptide of collagen type III
mannose/N-acetylglucosamine
complement
asialoglycoprotein
tissue plasminogen activator
low density lipoprotein
insulin
ceruloplasmin
enterokinase
carcinoembryonic antigen
apamin
galactose/lactose

Growth Factor/Cytokine receptors
hepatocyte growth factor
epidermal growth factor
insulin-like growth factor I, II
interleukin-1a/b
interleukin-2, IL-7, IL-4
γ-interferon
β-interferon
keratinocyte growth factor
TNF-R p55

Hormone receptors
prolactin
thyroglobulin
growth hormone
insulin
glucagon
leutinizing hormone
human choriogonadotrophic hormone

Nerve cell receptors
neurotensin

Antigen presenting cell receptors
immunoglobulin G-Fc receptor

Kidney cells
angiotensin II
vasopressin

Bone marrow receptors
c kit
CD-34

Keratinocyte and skin fibroblast receptors
very low density lipoprotein
low density lipoprotein
integrins that bind to RGD bearing polypeptides
collagen
laminin

Placental receptors
hemopexin
immunoglobulin G-Fc
low density lipoprotein
transferrin
alpha2-macroglobulin
ferritin
insulin
γ-interferon
epidermal growth factor
insulin-like growth factor

Muscle cell receptors
insulin
very low density lipoprotein

Gut epithelium
cobalamin-intrinsic factor
heat stable enterotoxin of E. Coli

Breast epithelium
heregulin
prolactin

Melanocytes
c kit

Miscellaneous
folate
cobalamin (B12)

Preferred ligands
low density lipoprotein (apoprotein B100)
very low density lipoprotein (apoprotein E)
galactose
c kit ligand
transferrin
insulin
heregulin
RGD or RGD-containing polypeptides

Table 7

Exemplary Antigens to Induce or Enhance an Immune Response

Melanoma or other tumor specific antigens;
leishmaniasis antigens;
helicobacter pylori specific antigens (e.g., urease B);
hepatitis B antigens;
hepatitis C antigens;
Herpes simplex antigens;
HIV antigens;
Tuberculosis antigens;
cytomegalovirus antigens;
lyme disease antigens;
malaria antigens;
respiratory syncytial virus antigens;
leprosy antigens;
toxoplasmosis antigens;
pneumocytis carinii antigens;
schistosomiasis antigens;
chlamydial antigens;
HTLV-1 antigens;
enterococcal antigens (e.g., VRE);
gonococcal antigens;
treponemal antigens;
clostridium difficile antigens;
Staphylococcus aureus antigens (e.g., MRSA);
trypanosomal antigens;
filarial antigens;
salmonella antigens;
shigella antigens;
pneumococcal antigens (e.g., penicillin resistant strains);
pseudomonal antigens

Table 8

Advantages of the Targeted Phage Vectors Over the Vectors Currently Used in Gene Therapy

	Size Constraints	**Specificity of Targeting**	**Immunogenicity/ Toxicity**	**Sustained/High/Low/ Controlled Expression**
Retrovirus Vector	7Kb	none	none	low, uncontrolled transient transfection
Adenovirus Vector	7Kb	none	high immunogenicity	low, uncontrolled transient transfection
Liposome	none	none	toxic at high doses	low, uncontrolled transient transfection
Phage Vectors	bacteriophage lambda DASH II (up to 23Kb); Minimal lambda cassette (up to 50 kb); P1 (up to 95 Kb)	yes	none	controlled high persistent expression

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: BRIGHAM AND WOMEN'S HOSPITAL, INC.
(B) STREET: 75 FRANCIS STREET
(C) CITY: BOSTON
(D) STATE: MASSACHUSETTS
(E) COUNTRY: UNITED STATES OF AMERICA
(F) POSTAL CODE: 02115

(i) APPLICANT/INVENTOR:
(A) NAME: SARKAR, SAUMYENDRA N.
(B) STREET: 102 QUEENSBURY STREET
(C) CITY: BOSTON
(D) STATE: MASSACHUSETTS
(E) COUNTRY: UNITED STATES OF AMERICA
(F) POSTAL CODE: 02115

(i) APPLICANT/INVENTOR:
(A) NAME: KUPPER, THOMAS S.
(B) STREET: 8 SURREY LANE
(C) CITY: WESTON
(D) STATE: MASSACHUSETTS
(E) COUNTRY: UNITED STATES OF AMERICA
(F) POSTAL CODE: 02193

(i) APPLICANT/INVENTOR:
(A) NAME: DUBIN, DANIEL B.
(B) STREET: 99 POND AVE
(C) CITY: BROOKLINE
(D) STATE: MASSACHUSETTS
(E) COUNTRY: UNITED STATES OF AMERICA
(F) POSTAL CODE: 02146

(ii) TITLE OF THE INVENTION: BACTERIOPHAGE-MEDIATED GENE THERAPY

(iii) NUMBER OF SEQUENCES: 21

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
(B) STREET: 600 ATLANTIC AVENUE
(C) CITY: BOSTON
(D) STATE: MASSACHUSETTS

(E) COUNTRY: UNITED STATES OF AMERICA
(F) POSTAL CODE: 02210

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 08/693,865
(B) FILING DATE: 08-MAY-1996

(A) APPLICATION NUMBER: US 08/814,859
(B) FILING DATE: 11-MAR-1997

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Plumer, Elizabeth R.
(B) REGISTRATION NUMBER: 36,637
(C) REFERENCE/DOCKET NUMBER: B0801/7059WO

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 617-720-3500
(B) TELEFAX: 617-720-2441

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Bacteriophage lambda

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATACCGAGGG CTGCAGTGTA CA 22

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 41 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Bacteriophage lambda

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

CTCTTTCAAT TGGGGAGGCA AAACGATGCT GATTGCCGTT C 41

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

TTGCCTCCCC AATTGAAAGA G 21

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 60 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GTGATGAAGG GTAAAGTTAT TTGCGTTTTT TTTTCGGCGG GGTCCTCCAT AAATTCAATC 60

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Bacteriophage lambda

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

TAACTTTACC CTTCATCACT AAAGGCC 27

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Bacteriophage lambda

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

AAACGTACAG CGCCATGTTT ACCAG 25

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:
Pro Lys Lys Lys Arg Lys Val
1 5

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Val Ser Arg Lys Arg Pro Arg Pro
1 5

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Val Arg Thr Thr Lys Gly Lys Arg Lys Arg Ile Asp Val
1 5 10

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Ala Ala Lys Arg Val Lys Leu Asp
1 5

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:
Leu Ser Ser Lys Arg Pro Arg Pro
1 5

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Lys Leu Lys His
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 300 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met Thr Ser Arg Arg Ser Val Lys Ser Gly Pro His Glu Val Pro Arg
1 5 10 15

Gln Glu Tyr Glu Asp Leu Tyr Tyr Thr Pro Ser Ser Gly Met Ala Ser
20 25 30

Pro Gln Ser Pro Pro Gln Thr Ser Arg Arg Gly Ala Leu Gln Thr Arg
35 40 45

Ser Arg Gln Arg Gly Glu Val Arg Phe Val Gly Tyr Asp Glu Ser Asp
50 55 60

Tyr Ala Leu Tyr Gly Gly Ser Ser Ser Glu Asp Asp Glu His Pro Glu
65 70 75 80

Pro Pro Thr Arg Arg Pro Val Ser Gly Ala Val Ala Ser Gly Pro Gly
85 90 95

Pro Ala His Ala Pro Pro Pro Pro Ala Gly Ser Gly Gly Ala Gly Arg
100 105 110

Thr Pro Thr Thr Ala Pro Arg Ala Pro Arg Thr Gln Arg Val Ala Thr
115 120 125

Lys Ala Pro Ala Ala Pro Ala Ala Glu Thr Thr Arg Gly Arg Lys Ser
130 135 140

Ala Gln Pro Glu Ser Ala Ala Leu Pro Gln Ala Pro Ala Ser Thr Ala
145 150 155 160

Arg Thr Arg Ser Lys Thr Pro Ala Gly Gly Leu Ala Arg Lys Leu His
165 170 175

Glu Ser Thr Ala Pro Pro Asn Pro Asp Ala Pro Val Val Thr Pro Arg
180 185 190

Val Ala Gly Phe Asn Lys Arg Val Cys Ala Ala Val Gly Arg Leu Ala
195 200 205

Ala Met His Ala Arg Met Ala Ala Val Gln Leu Val Val Asp Met Ser
210 215 220

Arg Pro Arg Ile Asp Glu Asp Ile Asn Glu Leu Leu Gly Ile Thr Thr
225 230 235 240

Ile Arg Val Thr Val Cys Glu Gly Lys Asn Leu Leu Gln Arg Ala Asn
245 250 255

Glu Val Asn Pro Asp Val Val Gln Asp Val Asp Ala Ala Thr Ala Thr
260 265 270

Arg Gly Arg Ser Ala Ala Ser Arg Pro Thr Glu Arg Pro Arg Ala Pro
275 280 285

Ala Pro Ser Ala Ser Arg Pro Arg Arg Pro Val Glu
290 295 300

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9009 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CAGTGGGTAA GTTCCTTAAA TGCTCTGCAA AGAAATTGGG ACTTTTCATT AAATCAGAAA 60
TTTTACTTTT TTCCCCTCCT GGGAGCTAAA GATATTTTAG AGAAGAATTA ACCTTTTGCT 120
TCTCCAGTTG AACATTTGTA GCAATAAGTC ATGCAAATAG AGCTCTCCAC CTGCTTCTTT 180
CTGTGCCTTT TGCGATTCTG CTTTAGTGCC ACCAGAAGAT ACTACCTGGG TGCAGTGGAA 240
CTGTCATGGG ACTATATGCA AAGTGATCTC GGTGAGCTGC CTGTGGACGC AAGATTTCCT 300
CCTAGAGTGC CAAAATCTTT TCCATTCAAC ACCTCAGTCG TGTACAAAAA GACTCTGTTT 360
GTAGAATTCA CGGTTCACCT TTTCAACATC GCTAAGCCAA GGCCACCCTG GATGGGTCTG 420
CTAGGTCCTA CCATCCAGGC TGAGGTTTAT GATACAGTGG TCATTACACT TAAGAACATG 480
GCTTCCCATC CTGTCAGTCT TCATGCTGTT GGTGTATCCT ACTGGAAAGC TTCTGAGGGA 540
GCTGAATATG ATGATCAGAC CAGTCAAAGG GAGAAAGAAG ATGATAAAGT CTTCCCTGGT 600
GGAAGCCATA CATATGTCTG GCAGGTCCTG AAAGAGAATG GTCCAATGGC CTCTGACCCA 660
CTGTGCCTTA CCTACTCATA TCTTTCTCAT GTGGACCTGG TAAAAGACTT GAATTCAGGC 720
CTCATTGGAG CCCTACTAGT ATGTAGAGAA GGGAGTCTGG CCAAGGAAAA GACACAGACC 780
TTGCACAAAT TTATACTACT TTTTGCTGTA TTTGATGAAG GGAAAAGTTG GCACTCAGAA 840
ACAAAGAACT CCTTGATGCA GGATAGGGAT GCTGCATCTG CTCGGGCCTG GCCTAAAATG 900
CACACAGTCA ATGGTTATGT AAACAGGTCT CTGCCAGGTC TGATTGGATG CCACAGGAAA 960
TCAGTCTATT GGCATGTGAT TGGAATGGGC ACCACTCCTG AAGTGCACTC AATATTCCTC 1020
GAAGGTCACA CATTTCTTGT GAGGAACCAT CGCCAGGCGT CCTTGGAAAT CTCGCCAATA 1080
ACTTTCCTTA CTGCTCAAAC ACTCTTGATG GACCTTGGAC AGTTTCTACT GTTTTGTCAT 1140
ATCTCTTCCC ACCAACATGA TGGCATGGAA GCTTATGTCA AAGTAGACAG CTGTCCAGAG 1200
GAACCCCAAC TACGAATGAA AAATAATGAA GAAGCGGAAG ACTATGATGA TGATCTTACT 1260

GATTCTGAAA TGGATGTGGT CAGGTTTGAT GATGACAACT CTCCTTCCTT TATCCAAATT 1320
CGCTCAGTTG CCAAGAAGCA TCCTAAAACT TGGGTACATT ACATTGCTGC TGAAGAGGAG 1380
GACTGGGACT ATGCTCCCTT AGTCCTCGCC CCCGATGACA GAAGTTATAA AAGTCAATAT 1440
TTGAACAATG GCCCTCAGCG GATTGGTAGG AAGTACAAAA AAGTCCGATT TATGGCATAC 1500
ACAGATGAAA CCTTTAAGAC TCGTGAAGCT ATTCAGCATG AATCAGGAAT CTTGGGACCT 1560
TTACTTTATG GGGAAGTTGG AGACACACTG TTGATTATAT TTAAGAATCA AGCAAGCAGA 1620
CCATATAACA TCTACCCTCA CGGAATCACT GATGTCCGTC CTTTGTATTC AAGGAGATTA 1680
CCAAAAGGTG TAAAACATTT GAAGGATTTT CCAATTCTGC CAGGAGAAAT ATTCAAATAT 1740
AAATGGACAG TGACTGTAGA AGATGGGCCA ACTAAATCAG ATCCTCGGTG CCTGACCCGC 1800
TATTACTCTA GTTTCGTTAA TATGGAGAGA GATCTAGCTT CAGGACTCAT TGGCCCTCTC 1860
CTCATCTGCT ACAAAGAATC TGTAGATCAA AGAGGAAACC AGATAATGTC AGACAAGAGG 1920
AATGTCATCC TGTTTTCTGT ATTTGATGAG AACCGAAGCT GGTACCTCAC AGAGAATATA 1980
CAACGCTTTC TCCCCAATCC AGCTGGAGTG CAGCTTGAGG ATCCAGAGTT CCAAGCCTCC 2040
AACATCATGC ACAGCATCAA TGGCTATGTT TTTGATAGTT TGCAGTTGTC AGTTTGTTTG 2100
CATGAGGTGG CATACTGGTA CATTCTAAGC ATTGGAGCAC AGACTGACTT CCTTTCTGTC 2160
TTCTTCTCTG GATATACCTT CAAACACAAA ATGGTCTATG AAGACACACT CACCCTATTC 2220
CCATTCTCAG GAGAAACTGT CTTCATGTCG ATGGAAAACC CAGGTCTATG GATTCTGGGG 2280
TGCCACAACT CAGACTTTCG GAACAGAGGC ATGACCGCCT TACTGAAGGT TTCTAGTTGT 2340
GACAAGAACA CTGGTGATTA TTACGAGGAC AGTTATGAAG ATATTTCAGC ATACTTGCTG 2400
AGTAAAAACA ATGCCATTGA ACCAAGAAGC TTCTCCCAGA ATTCAAGACA CCCTAGCACT 2460
AGGCAAAAGC AATTTAATGC CACCACAATT CCAGAAAATG ACATAGAGAA GACTGACCCT 2520
TGGTTTGCAC ACAGAACACC TATGCCTAAA ATACAAAATG TCTCCTCTAG TGATTTGTTG 2580
ATGCTCTTGC GACAGAGTCC TACTCCACAT GGGCTATCCT TATCTGATCT CCAAGAAGCC 2640
AAATATGAGA CTTTTTCTGA TGATCCATCA CCTGGAGCAA TAGACAGTAA TAACAGCCTG 2700
TCTGAAATGA CACACTTCAG GCCACAGCTC CATCACAGTG GGGACATGGT ATTTACCCCT 2760
GAGTCAGGCC TCCAATTAAG ATTAAATGAG AAACTGGGGA CAACTGCAGC AACAGAGTTG 2820
AAGAAACTTG ATTTCAAAGT TTCTAGTACA TCAAATAATC TGATTTCAAC AATTCCATCA 2880
GACAATTTGG CAGCAGGTAC TGATAATACA AGTTCCTTAG GACCCCCAAG TATGCCAGTT 2940
CATTATGATA GTCAATTAGA TACCACTCTA TTTGGCAAAA AGTCATCTCC CCTTACTGAG 3000
TCTGGTGGAC CTCTGAGCTT GAGTGAAGAA AATAATGATT CAAAGTTGTT AGAATCAGGT 3060
TTAATGAATA GCCAAGAAAG TTCATGGGGA AAAAATGTAT CGTCAACAGA GAGTGGTAGG 3120
TTATTTAAAG GGAAAAGAGC TCATGGACCT GCTTTGTTGA CTAAAGATAA TGCCTTATTC 3180
AAAGTTAGCA TCTCTTTGTT AAAGACAAAC AAAACTTCCA ATAATTCAGC AACTAATAGA 3240
AAGACTCACA TTGATGGCCC ATCATTATTA ATTGAGAATA GTCCATCAGT CTGGCAAAAT 3300
ATATTAGAAA GTGACACTGA GTTTAAAAAA GTGACACCTT TGATTCATGA CAGAATGCTT 3360
ATGGACAAAA ATGCTACAGC TTTGAGGCTA AATCATATGT CAAATAAAAC TACTTCATCA 3420
AAAAACATGG AAATGGTCCA ACAGAAAAAA GAGGGCCCCA TTCCACCAGA TGCACAAAAT 3480
CCAGATATGT CGTTCTTTAA GATGCTATTC TTGCCAGAAT CAGCAAGGTG GATACAAAGG 3540
ACTCATGGAA AGAACTCTCT GAACTCTGGG CAAGGCCCCA GTCCAAAGCA ATTAGTATCC 3600
TTAGGACCAG AAAAATCTGT GGAAGGTCAG AATTTCTTGT CTGAGAAAAA CAAAGTGGTA 3660
GTAGGAAAGG GTGAATTTAC AAAGGACGTA GGACTCAAAG AGATGGTTTT TCCAAGCAGC 3720
AGAAACCTAT TTCTTACTAA CTTGGATAAT TTACATGAAA ATAATACACA CAATCAAGAA 3780
AAAAAAATTC AGGAAGAAAT AGAAAAGAAG GAAACATTAA TCCAAGAGAA TGTAGTTTTG 3840
CCTCAGATAC ATACAGTGAC TGGCACTAAG AATTTCATGA AGAACCTTTT CTTACTGAGC 3900
ACTAGGCAAA ATGTAGAAGG TTCATATGAG GGGGCATATG CTCCAGTACT TCAAGATTTT 3960
AGGTCATTAA ATGATTCAAC AAATAGAACA AAGAAACACA CAGCTCATTT CTCAAAAAAA 4020
GGGGAGGAAG AAAACTTGGA AGGCTTGGGA AATCAAACCA AGCAAATTGT AGAGAAATAT 4080

GCATGCACCA CAAGGATATC TCCTAATACA AGCCAGCAGA ATTTTGTCAC GCAACGTAGT 4140
AAGAGAGCTT TGAAACAATT CAGACTCCCA CTAGAAGAAA CAGAACTTGA AAAAAGGATA 4200
ATTGTGGATG ACACCTCAAC CCAGTGGTCC AAAAACATGA AACATTTGAC CCCGAGCACC 4260
CTCACACAGA TAGACTACAA TGAGAAGGAG AAAGGGGCCA TTACTCAGTC TCCCTTATCA 4320
GATTGCCTTA CGAGGAGTCA TAGCATCCCT CAAGCAAATA GATCTCCATT ACCCATTGCA 4380
AAGGTATCAT CATTTCCATC TATTAGACCT ATATATCTGA CCAGGGTCCT ATTCCAAGAC 4440
AACTCTTCTC ATCTTCCAGC AGCATCTTAT AGAAAGAAAG ATTCTGGGGT CCAAGAAAGC 4500
AGTCATTTCT TACAAGGAGC CAAAAAAAAT AACCTTTCTT TAGCCATTCT AACCTTGGAG 4560
ATGACTGGTG ATCAAAGAGA GGTTGGCTCC CTGGGGACAA GTGCCACAAA TTCAGTCACA 4620
TACAAGAAAG TTGAGAACAC TGTTCTCCCG AAACCAGACT TGCCCAAAAC ATCTGGCAAA 4680
GTTGAATTGC TTCCAAAAGT TCACATTTAT CAGAAGGACC TATTCCCTAC GGAAACTAGC 4740
AATGGGTCTC CTGGCCATCT GGATCTCGTG GAAGGGAGCC TTCTTCAGGG AACAGAGGGA 4800
GCGATTAAGT GGAATGAAGC AAACAGACCT GGAAAAGTTC CCTTTCTGAG AGTAGCAACA 4860
GAAAGCTCTG CAAAGACTCC CTCCAAGCTA TTGGATCCTC TTGCTTGGGA TAACCACTAT 4920
GGTACTCAGA TACCAAAAGA AGAGTGGAAA TCCCAAGAGA AGTCACCAGA AAAAACAGCT 4980
TTTAAGAAAA AGGATACCAT TTTGTCCCTG AACGCTTGTG AAAGCAATCA TGCAATAGCA 5040
GCAATAAATG AGGGACAAAA TAAGCCCGAA ATAGAAGTCA CCTGGGCAAA GCAAGGTAGG 5100
ACTGAAAGGC TGTGCTCTCA AAACCCACCA GTCTTGAAAC GCCATCAACG GGAAATAACT 5160
CGTACTACTC TTCAGTCAGA TCAAGAGGAA ATTGACTATG ATGATACCAT ATCAGTTGAA 5220
ATGAAGAAGG AAGATTTTGA CATTTATGAT GAGGATGAAA ATCAGAGCCC CCGCAGCTTT 5280
CAAAAGAAAA CACGACACTA TTTTATTGCT GCAGTGGAGA GGCTCTGGGA TTATGGGATG 5340
AGTAGCTCCC CACATGTTCT AAGAAACAGG GCTCAGAGTG GCAGTGTCCC TCAGTTCAAG 5400
AAAGTTGTTT TCCAGGAATT TACTGATGGC TCCTTTACTC AGCCCTTATA CCGTGGAGAA 5460
CTAAATGAAC ATTTGGGACT CCTGGGGCCA TATATAAGAG CAGAAGTTGA AGATAATATC 5520
ATGGTAACTT TCAGAAATCA GGCCTCTCGT CCCTATTCCT TCTATTCTAG CCTTATTTCT 5580
TATGAGGAAG ATCAGAGGCA AGGAGCAGAA CCTAGAAAAA ACTTTGTCAA GCCTAATGAA 5640
ACCAAAACTT ACTTTTGGAA AGTGCAACAT CATATGGCAC CCACTAAAGA TGAGTTTGAC 5700
TGCAAAGCCT GGGCTTATTT CTCTGATGTT GACCTGGAAA AAGATGTGCA CTCAGGCCTG 5760
ATTGGACCCC TTCTGGTCTG CCACACTAAC ACACTGAACC CTGCTCATGG GAGACAAGTG 5820
ACAGTACAGG AATTTGCTCT GTTTTTCACC ATCTTTGATG AGACCAAAAG CTGGTACTTC 5880
ACTGAAAATA TGGAAAGAAA CTGCAGGGCT CCCTGCAATA TCCAGATGGA AGATCCCACT 5940
TTTAAAGAGA ATTATCGCTT CCATGCAATC AATGGCTACA TAATGGATAC ACTACCTGGC 6000
TTAGTAATGG CTCAGGATCA AAGGATTCGA TGGTATCTGC TCAGCATGGG CAGCAATGAA 6060
AACATCCATT CTATTCATTT CAGTGGACAT GTGTTCACTG TACGAAAAAA AGAGGAGTAT 6120
AAAATGGCAC TGTACAATCT CTATCCAGGT GTTTTTGAGA CAGTGGAAAT GTTACCATCC 6180
AAAGCTGGAA TTTGGCGGGT GGAATGCCTT ATTGGCGAGC ATCTACATGC TGGGATGAGC 6240
ACACTTTTTC TGGTGTACAG CAATAAGTGT CAGACTCCCC TGGGAATGGC TTCTGGACAC 6300
ATTAGAGATT TTCAGATTAC AGCTTCAGGA CAATATGGAC AGTGGGCCCC AAAGCTGGCC 6360
AGACTTCATT ATTCCGGATC AATCAATGCC TGGAGCACCA AGGAGCCCTT TTCTTGGATC 6420
AAGGTGGATC TGTTGGCACC AATGATTATT CACGGCATCA AGACCCAGGG TGCCCGTCAG 6480
AAGTTCTCCA GCCTCTACAT CTCTCAGTTT ATCATCATGT ATAGTCTTGA TGGGAAGAAG 6540
TGGCAGACTT ATCGAGGAAA TTCCACTGGA ACCTTAATGG TCTTCTTTGG CAATGTGGAT 6600
TCATCTGGGA TAAAACACAA TATTTTTAAC CCTCCAATTA TTGCTCGATA CATCCGTTTG 6660
CACCCAACTC ATTATAGCAT TCGCAGCACT CTTCGCATGG AGTTGATGGG CTGTGATTTA 6720
AATAGTTGCA GCATGCCATT GGGAATGGAG AGTAAAGCAA TATCAGATGC ACAGATTACT 6780
GCTTCATCCT ACTTTACCAA TATGTTTGCC ACCTGGTCTC CTTCAAAAGC TCGACTTCAC 6840
CTCCAAGGGA GGAGTAATGC CTGGAGACCT CAGGTGAATA ATCCAAAAGA GTGGCTGCAA 6900

GTGGACTTCC AGAAGACAAT GAAAGTCACA GGAGTAACTA CTCAGGGAGT AAAATCTCTG 6960
CTTACCAGCA TGTATGTGAA GGAGTTCCTC ATCTCCAGCA GTCAAGATGG CCATCAGTGG 7020
ACTCTCTTTT TTCAGAATGG CAAAGTAAAG GTTTTTCAGG GAAATCAAGA CTCCTTCACA 7080
CCTGTGGTGA ACTCTCTAGA CCCACCGTTA CTGACTCGCT ACCTTCGAAT TCACCCCCAG 7140
AGTTGGGTGC ACCAGATTGC CCTGAGGATG GAGGTTCTGG GCTGCGAGGC ACAGGACCTC 7200
TACTGAGGGT GGCCACTGCA GCACCTGCCA CTGCCGTCAC CTCTCCCTCC TCAGCTCCAG 7260
GGCAGTGTCC CTCCCTGGCT TGCCTTCTAC CTTTGTGCTA AATCCTAGCA GACACTGCCT 7320
TGAAGCCTCC TGAATTAACT ATCATCAGTC CTGCATTTCT TTGGTGGGGG GCCAGGAGGG 7380
TGCATCCAAT TTAACTTAAC TCTTACCTAT TTTCTGCAGC TGCTCCCAGA TTACTCCTTC 7440
CTTCCAATAT AACTAGGCAA AAAGAAGTGA GGAGAAACCT GCATGAAAGC ATTCTTCCCT 7500
GAAAAGTTAG GCCTCTCAGA GTCACCACTT CCTCTGTTGT AGAAAAACTA TGTGATGAAA 7560
CTTTGAAAAA GATATTTATG ATGTTAACAT TTCAGGTTAA GCCTCATACG TTTAAAATAA 7620
AACTCTCAGT TGTTTATTAT CCTGATCAAG CATGGAACAA AGCATGTTTC AGGATCAGAT 7680
CAATACAATC TTGGAGTCAA AAGGCAAATC ATTTGGACAA TCTGCAAAAT GGAGAGAATA 7740
CAATAACTAC TACAGTAAAG TCTGTTTCTG CTTCCTTACA CATAGATATA ATTATGTTAT 7800
TTAGTCATTA TGAGGGGCAC ATTCTTATCT CCAAAACTAG CATTCTTAAA CTGAGAATTA 7860
TAGATGGGGT TCAAGAATCC CTAAGTCCCC TGAAATTATA TAAGGCATTC TGTATAAATG 7920
CAAATGTGCA TTTTTCTGAC GAGTGTCCAT AGATATAAAG CCATTGGTCT TAATTCTGAC 7980
CAATAAAAAA ATAAGTCAGG AGGATGCAAT TGTTGAAAGC TTTGAAATAA AATAACATGT 8040
CTTCTTGAAA TTTGTGATGG CCAAGAAAGA AAATGATGAT GACATTAGGC TTCTAAAGGA 8100
CATACATTTA ATATTTCTGT GGAAATATGA GGAAAATCCA TGGTTATCTG AGATAGGAGA 8160
TACAAACTTT GTAATTCTAA TAATGCACTC AGTTTACTCT CTCCCTCTAC TAATTTCCTG 8220
CTGAAAATAA CACAACAAAA ATGTAACAGG GGAAATTATA TACCGTGACT GAAAACTAGA 8280
GTCCTACTTA CATAGTTGAA ATATCAAGGA GGTCAGAAGA AAATTGGACT GGTGAAAACA 8340
GAAAAAACAC TCCAGTCTGC CATATCACCA CACAATAGGA TCCCCCTTCT TGCCCTCCAC 8400
CCCCATAAGA TTGTGAAGGG TTTACTGCTC CTTCCATCTG CCTGCACCCC TTCACTATGA 8460
CTACACAGAA CTCTCCTGAT AGTAAAGGGG GCTGGAGGCA AGGATAAGTT ATAGAGCAGT 8520
TGGAGGAAGC ATCCAAAGAC TGCAACCCAG GGCAAATGGA AAACAGGAGA TCCTAATATG 8580
AAAGAAAAAT GGATCCCAAT CTGAGAAAAG GCAAAAGAAT GGCTACTTTT TTCTATGCTG 8640
GAGTATTTTC TAATAATCCT GCTTGACCCT TATCTGACCT CTTTGGAAAC TATAACATAG 8700
CTGTCACAGT ATAGTCACAA TCCACAAATG ATGCAGGTGC AAATGGTTTA TAGCCCTGTG 8760
AAGTTCTTAA AGTTTAGAGG CTAACTTACA GAAATGAATA AGTTGTTTTG TTTTATAGCC 8820
CGGTAGAGGA GTTAACCCCA AAGGTGATAT GGTTTTATTT CCTGTTATGT TTAACTTGAT 8880
AATCTTATTT TGGCATTCTT TTCCCATTGA CTATATACAT CTCTATTTCT CAAATGTTCA 8940
TGGAACTAGC TCTTTTATTT TCCTGCTGGT TTCTTCAGTA ATGAGTTAAA TAAAACATTG 9000
ACACATACA 9009

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8575 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GCAGCTGAGA GCATGGCCTA GGGTGGGCGG CACCATTGTC CAGCAGCTGA GTTTCCCAGG 60
GACCTTGGAG ATAGCCGCAG CCCTCATTTG CAGGGGAAGA TGATTCCTGC CAGATTTGCC 120
GGGGTGCTGC TTGCTCTGGC CCTCATTTTG CCAGGGACCC TTTGTGCAGA AGGAACTCGC 180
GGCAGGTCAT CCACGGCCCG ATGCAGCCTT TTCGGAAGTG ACTTCGTCAA CACCTTTGAT 240
GGGAGCATGT ACAGCTTTGC GGGATACTGC AGTTACCTCC TGGCAGGGGG CTGCCAGAAA 300
CGCTCCTTCT CGATTATTGG GGACTTCCAG AATGGCAAGA GAGTGAGCCT CTCCGTGTAT 360
CTTGGGGAAT TTTTTGACAT CCATTTGTTT GTCAATGGTA CCGTGACACA GGGGGACCAA 420
AGAGTCTCCA TGCCCTATGC CTCCAAAGGG CTGTATCTAG AAACTGAGGC TGGGTACTAC 480
AAGCTGTCCG GTGAGGCCTA TGGCTTTGTG GCCAGGATCG ATGGCAGCGG CAACTTTCAA 540
GTCCTGCTGT CAGACAGATA CTTCAACAAG ACCTGCGGGC TGTGTGGCAA CTTTAACATC 600
TTTGCTGAAG ATGACTTTAT GACCCAAGAA GGGACCTTGA CCTCGGACCC TTATGACTTT 660
GCCAACTCAT GGGCTCTGAG CAGTGGAGAA CAGTGGTGTG AACGGGCATC TCCTCCCAGC 720
AGCTCATGCA ACATCTCCTC TGGGGAAATG CAGAAGGGCC TGTGGGAGCA GTGCCAGCTT 780
CTGAAGAGCA CCTCGGTGTT TGCCCGCTGC CACCCTCTGG TGGACCCCGA GCCTTTTGTG 840
GCCCTGTGTG AGAAGACTTT GTGTGAGTGT GCTGGGGGGC TGGAGTGCGC CTGCCCTGCC 900
CTCCTGGAGT ACGCCCGGAC CTGTGCCCAG GAGGGAATGG TGCTGTACGG CTGGACCGAC 960
CACAGCGCGT GCAGCCCAGT GTGCCCTGCT GGTATGGAGT ATAGGCAGTG TGTGTCCCCT 1020
TGCGCCAGGA CCTGCCAGAG CCTGCACATC AATGAAATGT GTCAGGAGCG ATGCGTGGAT 1080
GGCTGCAGCT GCCCTGAGGG ACAGCTCCTG GATGAAGGCC TCTGCGTGGA GAGCACCGAG 1140
TGTCCCTGCG TGCATTCCGG AAAGCGCTAC CCTCCCGGCA CCTCCCTCTC TCGAGACTGC 1200
AACACCTGCA TTTGCCGAAA CAGCCAGTGG ATCTGCAGCA ATGAAGAATG TCCAGGGGAG 1260
TGCCTTGTCA CAGGTCAATC ACACTTCAAG AGCTTTGACA ACAGATACTT CACCTTCAGT 1320
GGGATCTGCC AGTACCTGCT GGCCCGGGAT TGCCAGGACC ACTCCTTCTC CATTGTCATT 1380
GAGACTGTCC AGTGTGCTGA TGACCGCGAC GCTGTGTGCA CCCGCTCCGT CACCGTCCGG 1440
CTGCCTGGCC TGCACAACAG CCTTGTGAAA CTGAAGCATG GGGCAGGAGT TGCCATGGAT 1500
GGCCAGGACG TCCAGCTCCC CCTCCTGAAA GGTGACCTCC GCATCCAGCA TACAGTGACG 1560
GCCTCCGTGC GCCTCAGCTA CGGGGAGGAC CTGCAGATGG ACTGGGATGG CCGCGGGAGG 1620
CTGCTGGTGA AGCTGTCCCC CGTCTATGCC GGGAAGACCT GCGGCCTGTG TGGGAATTAC 1680
AATGGCAACC AGGGCGACGA CTTCCTTACC CCCTCTGGGC TGGCGGAGCC CCGGGTGGAG 1740
GACTTCGGGA ACGCCTGGAA GCTGCACGGG GACTGCCAGG ACCTGCAGAA GCAGCACAGC 1800
GATCCCTGCG CCCTCAACCC GCGCATGACC AGGTTCTCCG AGGAGGCGTG CGCGGTCCTG 1860
ACGTCCCCCA CATTCGAGGC CTGCCATCGT GCCGTCAGCC CGCTGCCCTA CCTGCGGAAC 1920
TGCCGCTACG ACGTGTGCTC CTGCTCGGAC GGCCGCGAGT GCCTGTGCGG CGCCCTGGCC 1980
AGCTATGCCG CGGCCTGCGC GGGGAGAGGC GTGCGCGTCG CGTGGCGCGA GCCAGGCCGC 2040
TGTGAGCTGA ACTGCCCGAA AGGCCAGGTG TACCTGCAGT GCGGGACCCC CTGCAACCTG 2100
ACCTGCCGCT CTCTCTCTTA CCCGGATGAG GAATGCAATG AGGCCTGCCT GGAGGGCTGC 2160
TTCTGCCCCC CAGGGCTCTA CATGGATGAG AGGGGGGACT GCGTGCCCAA GGCCCAGTGC 2220
CCCTGTTACT ATGACGGTGA GATCTTCCAG CCAGAAGACA TCTTCTCAGA CCATCACACC 2280
ATGTGCTACT GTGAGGATGG CTTCATGCAC TGTACCATGA GTGGAGTCCC CGGAAGCTTG 2340
CTGCCTGACG CTGTCCTCAG CAGTCCCCTG TCTCATCGCA GCAAAAGGAG CCTATCCTGT 2400
CGGCCCCCCA TGGTCAAGCT GGTGTGTCCC GCTGACAACC TGCGGGCTGA AGGGCTCGAG 2460
TGTACCAAAA CGTGCCAGAA CTATGACCTG GAGTGCATGA GCATGGGCTG TGTCTCTGGC 2520
TGCCTCTGCC CCCCGGGCAT GGTCCGGCAT GAGAACAGAT GTGTGGCCCT GGAAAGGTGT 2580
CCCTGCTTCC ATCAGGGCAA GGAGTATGCC CCTGGAGAAA CAGTGAAGAT TGGCTGCAAC 2640

ACTTGTGTCT GTCGGGACCG GAAGTGGAAC TGCACAGACC ATGTGTGTGA TGCCACGTGC 2700
TCCACGATCG GCATGGCCCA CTACCTCACC TTCGACGGGC TCAAATACCT GTTCCCCGGG 2760
GAGTGCCAGT ACGTTCTGGT GCAGGATTAC TGCGGCAGTA ACCCTGGGAC CTTTCGGATC 2820
CTAGTGGGGA ATAAGGGATG CAGCCACCCC TCAGTGAAAT GCAAGAAACG GGTCACCATC 2880
CTGGTGGAGG GAGGAGAGAT TGAGCTGTTT GACGGGGAGG TGAATGTGAA GAGGCCCATG 2940
AAGGATGAGA CTCACTTTGA GGTGGTGGAG TCTGGCCGGT ACATCATTCT GCTGCTGGGC 3000
AAAGCCCTCT CCGTGGTCTG GGACCGCCAC CTGAGCATCT CCGTGGTCCT GAAGCAGACA 3060
TACCAGGAGA AAGTGTGTGG CCTGTGTGGG AATTTTGATG GCATCCAGAA CAATGACCTC 3120
ACCAGCAGCA ACCTCCAAGT GGAGGAAGAC CCTGTGGACT TTGGGAACTC CTGGAAAGTG 3180
AGCTCGCAGT GTGCTGACAC CAGAAAAGTG CCTCTGGACT CATCCCCTGC CACCTGCCAT 3240
AACAACATCA TGAAGCAGAC GATGGTGGAT TCCTCCTGTA GAATCCTTAC CAGTGACGTC 3300
TTCCAGGACT GCAACAAGCT GGTGGACCCC GAGCCATATC TGGATGTCTG CATTTACGAC 3360
ACCTGCTCCT GTGAGTCCAT TGGGGACTGC GCCTGCTTCT GCGACACCAT TGCTGCCTAT 3420
GCCCACGTGT GTGCCCAGCA TGGCAAGGTG GTGACCTGGA GGACGGCCAC ATTGTGCCCC 3480
CAGAGCTGCG AGGAGAGGAA TCTCCGGGAG AACGGGTATG AGTGTGAGTG GCGCTATAAC 3540
AGCTGTGCAC CTGCCTGTCA AGTCACGTGT CAGCACCCTG AGCCACTGGC CTGCCCTGTG 3600
CAGTGTGTGG AGGGCTGCCA TGCCCACTGC CCTCCAGGGA AAATCCTGGA TGAGCTTTTG 3660
CAGACCTGCG TTGACCCTGA AGACTGTCCA GTGTGTGAGG TGGCTGGCCG GCGTTTTGCC 3720
TCAGGAAAGA AAGTCACCTT GAATCCCAGT GACCCTGAGC ACTGCCAGAT TTGCCACTGT 3780
GATGTTGTCA ACCTCACCTG TGAAGCCTGC CAGGAGCCGG GAGGCCTGGT GGTGCCTCCC 3840
ACAGATGCCC CGGTGAGCCC CACCACTCTG TATGTGGAGG ACATCTCGGA ACCGCCGTTG 3900
CACGATTTCT ACTGCAGCAG GCTACTGGAC CTGGTCTTCC TGCTGGATGG CTCCTCCAGG 3960
CTGTCCGAGG CTGAGTTTGA AGTGCTGAAG GCCTTTGTGG TGGACATGAT GGAGCGGCTG 4020
CGCATCTCCC AGAAGTGGGT CCGCGTGGCC GTGGTGGAGT ACCACGACGG CTCCCACGCC 4080
TACATCGGGC TCAAGGACCG GAAGCGACCG TCAGAGCTGC GGCGCATTGC CAGCCAGGTG 4140
AAGTATGCGG GCAGCCAGGT GGCCTCCACC AGCGAGGTCT TGAAATACAC ACTGTTCCAA 4200
ATCTTCAGCA AGATCGACCG CCCTGAAGCC TCCCGCATCG CCCTGCTCCT GATGGCCAGC 4260
CAGGAGCCCC AACGGATGTC CCGGAACTTT GTCCGCTACG TCCAGGGCCT GAAGAAGAAG 4320
AAGGTCATTG TGATCCCGGT GGGCATTGGG CCCCATGCCA ACCTCAAGCA GATCCGCCTC 4380
ATCGAGAAGC AGGCCCCTGA GAACAAGGCC TTCGTGCTGA GCAGTGTGGA TGAGCTGGAG 4440
CAGCAAAGGG ACGAGATCGT TAGCTACCTC TGTGACCTTG CCCCTGAAGC CCCTCCTCCT 4500
ACTCTGCCCC CCCACATGGC ACAAGTCACT GTGGGCCCGG GGCTCTTGGG GGTTTCGACC 4560
CTGGGGCCCA AGAGGAACTC CATGGTTCTG GATGTGGCGT TCGTCCTGGA AGGATCGGAC 4620
AAAATTGGTG AAGCCGACTT CAACAGGAGC AAGGAGTTCA TGGAGGAGGT GATTCAGCGG 4680
ATGGATGTGG GCCAGGACAG CATCCACGTC ACGGTGCTGC AGTACTCCTA CATGGTGACC 4740
GTGGAGTACC CCTTCAGCGA GGCACAGTCC AAAGGGGACA TCCTGCAGCG GGTGCGAGAG 4800
ATCCGCTACC AGGGCGGCAA CAGGACCAAC ACTGGGCTGG CCCTGCGGTA CCTCTCTGAC 4860
CACAGCTTCT TGGTCAGCCA GGGTGACCGG GAGCAGGCGC CCAACCTGGT CTACATGGTC 4920
ACCGGAAATC CTGCCTCTGA TGAGATCAAG AGGCTGCCTG GAGACATCCA GGTGGTGCCC 4980
ATTGGAGTGG GCCCTAATGC CAACGTGCAG GAGCTGGAGA GGATTGGCTG GCCCAATGCC 5040
CCTATCCTCA TCCAGGACTT TGAGACGCTC CCCCGAGAGG CTCCTGACCT GGTGCTGCAG 5100
AGGTGCTGCT CCGGAGAGGG GCTGCAGATC CCCACCCTCT CCCCTGCACC TGACTGCAGC 5160
CAGCCCCTGG ACGTGATCCT TCTCCTGGAT GGCTCCTCCA GTTTCCCAGC TTCTTATTTT 5220
GATGAAATGA AGAGTTTCGC CAAGGCTTTC ATTTCAAAAG CCAATATAGG GCCTCGTCTC 5280
ACTCAGGTGT CAGTGCTGCA GTATGGAAGC ATCACCACCA TTGACGTGCC ATGGAACGTG 5340
GTCCCGGAGA AAGCCCATTT GCTGAGCCTT GTGGACGTCA TGCAGCGGGA GGGAGGCCCC 5400
AGCCAAATCG GGGATGCCTT GGGCTTTGCT GTGCGATACT TGACTTCAGA AATGCATGGT 5460

GCCAGGCCGG GAGCCTCAAA GGCGGTGGTC ATCCTGGTCA CGGACGTCTC TGTGGATTCA 5520
GTGGATGCAG CAGCTGATGC CGCCAGGTCC AACAGAGTGA CAGTGTTCCC TATTGGAATT 5580
GGAGATCGCT ACGATGCAGC CCAGCTACGG ATCTTGGCAG GCCCAGCAGG CGACTCCAAC 5640
GTGGTGAAGC TCCAGCGAAT CGAAGACCTC CCTACCATGG TCACCTTGGG CAATTCCTTC 5700
CTCCACAAAC TGTGCTCTGG ATTTGTTAGG ATTTGCATGG ATGAGGATGG GAATGAGAAG 5760
AGGCCCGGGG ACGTCTGGAC CTTGCCAGAC CAGTGCCACA CCGTGACTTG CCAGCCAGAT 5820
GGCCAGACCT TGCTGAAGAC TCATCGGGTC AACTGTGACC GGGGGCTGAG GCCTTCGTGC 5880
CCTAACAGCC AGTCCCCTGT TAAAGTGGAA GAGACCTGTG GCTGCCGCTG GACCTGCCCC 5940
TGCGTGTGCA CAGGCAGCTC CACTCGGCAC ATCGTGACCT TTGATGGGCA GAATTTCAAG 6000
CTGACTGGCA GCTGTTCTTA TGTCCTATTT CAAAACAAGG AGCAGGACCT GGAGGTGATT 6060
CTCCATAATG GTGCCTGCAG CCCTGGAGCA AGGCAGGGCT GCATGAAATC CATCGAGGTG 6120
AAGCACAGTG CCCTCTCCGT CGAGCTGCAC AGTGACATGG AGGTGACGGT GAATGGGAGA 6180
CTGGTCTCTG TTCCTTACGT GGGTGGGAAC ATGGAAGTCA ACGTTTATGG TGCCATCATG 6240
CATGAGGTCA GATTCAATCA CCTTGGTCAC ATCTTCACAT TCACTCCACA AAACAATGAG 6300
TTCCAACTGC AGCTCAGCCC CAAGACTTTT GCTTCAAAGA CGTATGGTCT GTGTGGGATC 6360
TGTGATGAGA ACGGAGCCAA TGACTTCATG CTGAGGGATG GCACAGTCAC CACAGACTGG 6420
AAAACACTTG TTCAGGAATG GACTGTGCAG CGGCCAGGGC AGACGTGCCA GCCCATCCTG 6480
GAGGAGCAGT GTCTTGTCCC CGACAGCTCC CACTGCCAGG TCCTCCTCTT ACCACTGTTT 6540
GCTGAATGCC ACAAGGTCCT GGCTCCAGCC ACATTCTATG CCATCTGCCA GCAGGACAGT 6600
TGCCACCAGG AGCAAGTGTG TGAGGTGATC GCCTCTTATG CCCACCTCTG TCGGACCAAC 6660
GGGGTCTGCG TTGACTGGAG GACACCTGAT TTCTGTGCTA TGTCATGCCC ACCATCTCTG 6720
GTCTACAACC ACTGTGAGCA TGGCTGTCCC CGGCACTGTG ATGGCAACGT GAGCTCCTGT 6780
GGGGACCATC CCTCCGAAGG CTGTTTCTGC CCTCCAGATA AAGTCATGTT GGAAGGCAGC 6840
TGTGTCCCTG AAGAGGCCTG CACTCAGTGC ATTGGTGAGG ATGGAGTCCA GCACCAGTTC 6900
CTGGAAGCCT GGGTCCCGGA CCACCAGCCC TGTCAGATCT GCACATGCCT CAGCGGGCGG 6960
AAGGTCAACT GCACAACGCA GCCCTGCCCC ACGGCCAAAG CTCCCACGTG TGGCCTGTGT 7020
GAAGTAGCCC GCCTCCGCCA GAATGCAGAC CAGTGCTGCC CCGAGTATGA GTGTGTGTGT 7080
GACCCAGTGA GCTGTGACCT GCCCCCAGTG CCTCACTGTG AACGTGGCCT CCAGCCCACA 7140
CTGACCAACC CTGGCGAGTG CAGACCCAAC TTCACCTGCG CCTGCAGGAA GGAGGAGTGC 7200
AAAAGAGTGT CCCCACCCTC CTGCCCCCCG CACCGTTTGC CCACCCTTCG GAAGACCCAG 7260
TGCTGTGATG AGTATGAGTG TGCCTGCAAC TGTGTCAACT CCACAGTGAG CTGTCCCCTT 7320
GGGTACTTGG CCTCAACCGC CACCAATGAC TGTGGCTGTA CCACAACCAC CTGCCTTCCC 7380
GACAAGGTGT GTGTCCACCG AAGCACCATC TACCCTGTGG GCCAGTTCTG GGAGGAGGGC 7440
TGCGATGTGT GCACCTGCAC CGACATGGAG GATGCCGTGA TGGGCCTCCG CGTGGCCCAG 7500
TGCTCCCAGA AGCCCTGTGA GGACAGCTGT CGGTCGGGCT TCACTTACGT TCTGCATGAA 7560
GGCGAGTGCT GTGGAAGGTG CCTGCCATCT GCCTGTGAGG TGGTGACTGG CTCACCGCGG 7620
GGGGACTCCC AGTCTTCCTG GAAGAGTGTC GGCTCCCAGT GGGCCTCCCC GGAGAACCCC 7680
TGCCTCATCA ATGAGTGTGT CCGAGTGAAG GAGGAGGTCT TTATACAACA AAGGAACGTC 7740
TCCTGCCCCC AGCTGGAGGT CCCTGTCTGC CCCTCGGGCT TTCAGCTGAG CTGTAAGACC 7800
TCAGCGTGCT GCCCAAGCTG TCGCTGTGAG CGCATGGAGG CCTGCATGCT CAATGGCACT 7860
GTCATTGGGC CCGGGAAGAC TGTGATGATC GATGTGTGCA CGACCTGCCG CTGCATGGTG 7920
CAGGTGGGGG TCATCTCTGG ATTCAAGCTG GAGTGCAGGA AGACCACCTG CAACCCCTGC 7980
CCCCTGGGTT ACAAGGAAGA AAATAACACA GGTGAATGTT GTGGGAGATG TTTGCCTACG 8040
GCTTGCACCA TTCAGCTAAG AGGAGGACAG ATCATGACAC TGAAGCGTGA TGAGACGCTC 8100
CAGGATGGCT GTGATACTCA CTTCTGCAAG GTCAATGAGA GAGGAGAGTA CTTCTGGGAG 8160
AAGAGGGTCA CAGGCTGCCC ACCCTTTGAT GAACACAAGT GTCTGGCTGA GGGAGGTAAA 8220
ATTATGAAAA TTCCAGGCAC CTGCTGTGAC ACATGTGAGG AGCCTGAGTG CAACGACATC 8280

ACTGCCAGGC TGCAGTATGT CAAGGTGGGA AGCTGTAAGT CTGAAGTAGA GGTGGATATC 8340
CACTACTGCC AGGGCAAATG TGCCAGCAAA GCCATGTACT CCATTGACAT CAACGATGTG 8400
CAGGACCAGT GCTCCTGCTG CTCTCCGACA CGGACGGAGC CCATGCAGGT GGCCCTGCAC 8460
TGCACCAATG GCTCTGTTGT GTACCATGAG GTTCTCAATG CCATGGAGTG CAAATGCTCC 8520
CCCAGGAAGT GCAGCAAGTG AGGCTGCTGC AGCTGCATGG GTGCCTGCTG CTGCC 8575

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13815 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CCTCACTCAC TTGCCCCTTA CAGGACTCAG CTCTTGAAGG CAATAGCTTT ATAGAAAAAA 60
CGAATAGGAA GACTTGAAGT GCTATTTTTT TTTTTTTTTT TGTCAAGGCT GCTGAAGTTT 120
ATTGGCTTCT CATCGTACCT AAGCCTCCTG GAGCAATAAA ACTGGGAGAA ACTTTTACCA 180
AGATTTTTAT CCCTGCCTTG ATATATACTT TTTCTTCCAA ATGCTTTGGT GGGAAGAAGT 240
AGAGGACTGT TATGAAAGAG AAGATGTTCA AAAGAAAACA TTCACAAAAT GGATAAATGC 300
ACAATTTTCT AAGTTTGGAA AGCAACACAT AGACAACCTC TTCAGTGACC TGCAGGATGG 360
AAAACGCCTC CTAGACCTCT TGGAAGGCCT TACAGGGCAA AAACTGCCAA AAGAAAAGGG 420
ATCTACAAGA GTTCATGCCC TGAACAATGT CAACAAGGCA CTGCGGGTCT TACAGAAAAA 480
TAATGTTGAT TTAGTGAATA TAGGAAGCAC TGACATAGTG GATGGAAATC ATAAACTCAC 540
TCTTGGTTTG ATTTGGAATA TAATCCTCCA CTGGCAGGTC AAAAATGTGA TGAAAACTAT 600
CATGGCTGGA TTGCAGCAAA CCAACAGTGA AAAGATTCTT CTGAGCTGGG TTCGACAGTC 660
AACACGTAAT TATCCACAGG TTAACGTCAT CAACTTCACC TCTAGCTGGT CCGACGGGTT 720
GGCTTTGAAT GCTCTTATCC ATAGTCACAG GCCCGACCTG TTTGATTGGA ATAGTGTGGT 780
TTCACAGCAC TCAGCCACCC AAAGACTGGA ACATGCCTTC AACATTGCAA AATGCCAGTT 840
AGGCATAGAA AAACTTCTTG ATCCTGAAGA TGTTGCTACC ACTTATCCAG ACAAGAAGTC 900
CATCTTAATG TACATCACAT CACTCTTTCA AGTTTTGCCA CAACAAGTGA GCATTGAAGC 960
CATTCAAGAA GTGGAAATGT TGCCCAGGAC ATCTTCAAAA GTAACTAGAG AAGAACATTT 1020
TCAATTACAT CACCAGATGC ATTACTCTCA ACAGATCACA GTCAGTCTAG CACAGGGCTA 1080
TGAACAAACT TCTTCATCTC CTAAGCCTCG ATTCAAGAGT TATGCCTTCA CACAGGCTGC 1140
TTATGTTGCC ACCTCTGATT CCACACAGAG CCCCTATCCT TCACAGCATT TGGAAGCTCC 1200
CAGAGACAAG TCACTTGACA GTTCATTGAT GGAGACGGAA GTAAATCTGG ATAGTTACCA 1260
AACTGCTTTA GAAGAAGTAC TTTCATGGCT TCTTTCTGCC GAGGATACAT TGCGAGCACA 1320
AGGAGAGATT TCAAATGATG TTGAAGAAGT GAAAGAACAG TTTCATGCTC ATGAGGGATT 1380
CATGATGGAT CTGACATCTC ATCAAGGACT TGTTGGTAAT GTTCTACAGT TAGGAAGTCA 1440
ACTAGTTGGA AAAGGGAAAT TATCAGAAGA TGAAGAAGCT GAAGTGCAAG AACAAATGAA 1500
TCTCCTAAAT TCAAGATGGG AATGTCTCAG GGTAGCTAGC ATGGAAAAAC AAAGCAAATT 1560
ACACAAAGTT CTAATGGATC TCCAGAATCA GAAATTAAAA GAACTAGATG ACTGGTTAAC 1620

AAAAACTGAA GAGAGAACTA AGAAAATGGA GGAAGAGCCC TTTGGACCTG ATCTTGAAGA 1680
TCTAAAATGC CAAGTACAAC AACATAAGGT GCTTCAAGAA GATCTAGAAC AGGAGCAGGT 1740
CAGGGTCAAC TCGCTCACTC ACATGGTAGT AGTGGTTGAT GAATCCAGCG GTGATCATGC 1800
AACAGCTGCT TTGGAAGAAC AACTTAAGGT ACTGGGAGAT CGATGGGCAA ATATCTGCAG 1860
ATGGACTGAA GACCGCTGGA TTGTTTTACA AGATATTCTT CTAAAATGGC AGCATTTTAC 1920
TGAAGAACAG TGCCTTTTTA GTACATGGCT TTCAGAAAAA GAAGATGCAA TGAAGAACAT 1980
TCAGACAAGT GGCTTTAAAG ATCAAAATGA AATGATGTCA AGTCTTCACA AAATATCTAC 2040
TTTAAAAATA GATCTAGAAA AGAAAAAGCC AACCATGGAA AAACTAAGTT CACTCAATCA 2100
AGATCTACTT TCGGCACTGA AAAATAAGTC AGTGACTCAA AAGATGGAAA TCTGGATGGA 2160
AAACTTTGCA CAACGTTGGG ACAATTTAAC CCAAAAACTT GAAAAGAGTT CAGCACAAAT 2220
TTCACAGGCT GTCACCACCA CTCAACCATC CCTAACACAG ACAACTGTAA TGGAAACGGT 2280
AACTATGGTG ACCACAAGGG AACAAATCAT GGTAAAACAT GCCCAAGAGG AACTTCCACC 2340
ACCACCTCCT CAAAAGAAGA GGCAGATAAC TGTGGATTCT GAACTCAGGA AAAGGTTGGA 2400
TGTCGATATA ACTGAACTTC ACAGTTGGAT TACTCGTTCA GAAGCTGTAT TACAGAGTTC 2460
TGAATTTGCA GTCTATCGAA AAGAAGGCAA CATCTCAGAC TTGCAAGAAA AAGTCAATGC 2520
CATAGCACGA GAAAAAGCAG AGAAGTTCAG AAAACTGCAA GATGCCAGCA GATCAGCTCA 2580
GGCCCTGGTG GAACAGATGG CAAATGAGGG TGTTAATGCT GAAAGTATCA GACAAGCTTC 2640
AGAACAACTG AACAGCCGGT GGACAGAATT CTGCCAATTG CTGAGTGAGA GAGTTAACTG 2700
GCTAGAGTAT CAAACCAACA TCATTACCTT TTATAATCAG CTACAACAAT TGGAACAGAT 2760
GACAACTACT GCCGAAAACT TGTTGAAAAC CCAGTCTACC ACCCTATCAG AGCCAACAGC 2820
AATTAAAAGC CAGTTAAAAA TTTGTAAGGA TGAAGTCAAC AGATTGTCAG CTCTTCAGCC 2880
TCAAATTGAG CAATTAAAAA TTCAGAGTCT ACAACTGAAA GAAAAGGGAC AGGGGCCAAT 2940
GTTTCTGGAT GCAGACTTTG TGGCCTTTAC TAATCATTTT AACCACATCT TTGATGGTGT 3000
GAGGGCCAAA GAGAAAGAGC TACAGACAAT TTTTGACACT TTACCACCAA TGCGCTATCA 3060
GGAGACAATG AGTAGCATCA GGACGTGGAT CCAGCAGTCA GAAAGCAAAC TCTCTGTACC 3120
TTATCTTAGT GTTACTGAAT ATGAAATAAT GGAGGAGAGA CTCGGGAAAT TACAGGCTCT 3180
GCAAAGTTCT TTGAAAGAGC AACAAAATGG CTTCAACTAT CTGAGTGACA CTGTGAAGGA 3240
GATGGCCAAG AAAGCACCTT CAGAAATATG CCAGAAATAT CTGTCAGAAT TTGAAGAGAT 3300
TGAGGGGCAC TGGAAGAAAC TTTCCTCCCA GTTGGTGGAA AGCTGCCAAA AGCTAGAAGA 3360
ACATATGAAT AAACTTCGAA AATTTCAGAA TCACATAAAA ACCTTACAGA AATGGATGGC 3420
TGAAGTTGAT GTTTTCCTGA AAGAGGAATG GCCTGCCCTG GGGGATGCTG AAATCCTGAA 3480
AAAACAGCTC AAACAATGCA GACTTTTAGT TGGTGATATT CAAACAATTC AGCCCAGTTT 3540
AAATAGTGTT AATGAAGGTG GGCAGAAGAT AAAGAGTGAA GCTGAACTTG AGTTTGCATC 3600
CAGACTGGAG ACAGAACTTA GAGAGCTTAA CACTCAGTGG GATCACATAT GCCGCCAGGT 3660
CTACACCAGA AAGGAAGCCT TAAAGGCAGG TTTGGATAAA ACCGTAAGCC TCCAAAAAGA 3720
TCTATCAGAG ATGCATGAGT GGATGACACA AGCTGAAGAA GAATATCTAG AGAGAGATTT 3780
TGAATATAAA ACTCCAGATG AATTACAGAC TGCTGTTGAA GAAATGAAGA GAGCTAAAGA 3840
AGAGGCACTA CAAAAAGAAA CTAAAGTGAA ACTCCTTACT GAGACTGTAA ATAGTGTAAT 3900
AGCTCACGCT CCACCCTCAG CACAAGAGGC CTTAAAAAAG GAACTTGAAA CTCTGACCAC 3960
CAACTACCAA TGGCTGTGCA CCAGGCTGAA TGGAAAATGC AAAACTTTGG AAGAAGTTTG 4020
GGCATGTTGG CATGAGTTAT TGTCATATTT AGAGAAAGCA AACAAGTGGC TCAATGAAGT 4080
AGAATTGAAA CTTAAAACCA TGGAAAATGT TCCTGCAGGA CCTGAGGAAA TCACTGAAGT 4140
GCTAGAATCT CTTGAAAATC TGATGCATCA TTCAGAGGAG AACCCAAATC AGATTCGTCT 4200
ATTGGCACAG ACTCTTACAG ATGGAGGAGT CATGGATGAA CTGATCAATG AGGAGCTTGA 4260
GACGTTTAAT TCTCGTTGGA GGGAACTACA TGAAGAGGCT GTGAGGAAAC AAAAGTTGCT 4320
TGAACAGAGT ATCCAGTCTG CCCAGGAAAT TGAAAAGTCC TTGCACTTAA TTCAGGAGTC 4380
GCTTGAATTC ATTGACAAGC AGTTGGCAGC TTATATCACT GACAAGGTGG ATGCAGCTCA 4440

AATGCCTCAG GAAGCCCAGA AAATCCAATC AGATTTGACA AGTCATGAGA TAAGTTTAGA 4500
AGAAATGAAG AAACATAACC AGGGGAAGGA TGCCAACCAA AGGGTTCTTT CACAAATTGA 4560
TGTTGCACAG AAAAAATTAC AAGATGTCTC CATGAAATTT CGATTATTCC AAAAACCAGC 4620
CAATTTTGAA CAACGTCTAG AGGAAAGTAA GATGATTTTA GATGAAGTCA AGATGCATTT 4680
GCCTGCATTG GAAACCAAGA GTGTTGAACA GGAAGTAATT CAGTCACAAC TAAGTCATTG 4740
TGTGAACTTG TATAAAAGCC TGAGTGAAGT CAAGTCTGAA GTGGAAATGG TGATTAAAAC 4800
CGGACGTCAA ATTGTACAGA AAAAGCAGAC AGAAAATCCC AAAGAGCTTG ATGAACGAGT 4860
AACAGCTTTG AAATTGCATT ACAATGAGTT GGGTGCGAAG GTAACAGAGA GAAAGCAACA 4920
GTTGGAGAAA TGCTTGAAGT TGTCCCGTAA GATGAGAAAG GAAATGAATG TCTTAACAGA 4980
ATGGCTGGCA GCAACAGATA CAGAATTGAC GAAGAGATCA GCAGTTGAAG GAATGCCAAG 5040
TAATTTGGAT TCTGAAGTTG CCTGGGGAAA GGCTACTCAA AAAGAGATTG AGAAACAGAA 5100
GGCTCACTTG AAGAGTGTTA CAGAATTAGG AGAGTCTTTG AAAATGGTGT TGGGCAAGAA 5160
AGAAACCTTG GTAGAAGATA AACTGAGTCT TCTGAACAGT AACTGGATAG CTGTCACCTC 5220
CAGAGTAGAA GAATGGCTAA ATCTTTTGTT GGAATACCAG AAACACATGG AAACCTTTGA 5280
TCAGAACATA GAACAAATCA CAAAGTGGAT CATTCATGCA GATGAACTTT TAGATGAGTC 5340
TGAAAAGAAG AAACCACAAC AAAAGGAAGA CATTCTTAAG CGTTTAAAGG CTGAAATGAA 5400
TGACATGCGC CCAAAGGTGG ACTCCACACG TGACCAAGCA GCAAAATTGA TGGCAAACCG 5460
CGGTGACCAC TGCAGGAAAG TAGTAGAGCC CCAAATCTCT GAGCTCAACC GTCGATTTGC 5520
AGCTATTTCT CACAGAATTA AGACTGGAAA GGCCTCCATT CCTTTGAAGG AATTGGAGCA 5580
GTTTAACTCA GATATACAAA AATTGCTTGA ACCACTGGAG GCTGAAATTC AGCAGGGGGT 5640
GAATCTGAAA GAGGAAGACT TCAATAAAGA TATGAGTGAA GACAATGAGG GTACTGTAAA 5700
TGAATTGTTG CAAAGAGGAG ACAACTTACA ACAAAGAATC ACAGATGAGA GAAAGCGAGA 5760
GGAAATAAAG ATAAAACAGC AGCTGTTACA GACAAAACAT AATGCTCTCA AGGATTTGAG 5820
GTCTCAAAGA AGAAAAAAGG CCCTAGAAAT TTCTCACCAG TGGTATCAGT ACAAGAGGCA 5880
GGCTGATGAT CTCCTGAAAT GCTTGGATGA AATTGAAAAA AAATTAGCCA GCCTACCTGA 5940
ACCCAGAGAT GAAAGAAAAT TAAAGGAAAT TGATCGTGAA TTGCAGAAGA AGAAAGAGGA 6000
GCTGAATGCA GTGCGCAGGC AAGCTGAGGG CTTGTCTGAG AATGGGGCCG CAATGGCAGT 6060
GGAGCCAACT CAGATCCAGC TCAGCAAGCG CTGGCGGCAA ATTGAGAGCA ATTTTGCTCA 6120
GTTTCGAAGA CTCAACTTTG CACAAATTCA CACTCTCCAT GAAGAAACTA TGGTAGTGAC 6180
GACTGAAGAT ATGCCTTTGG ATGTTTCTTA TGTGCCTTCT ACTTATTTGA CCGAGATCAG 6240
TCATATCTTA CAAGCTCTTT CAGAAGTTGA TCATCTTCTA AATACTCCTG AACTCTGTGC 6300
TAAAGATTTT GAAGATCTTT TTAAGCAAGA GGAGTCTCTT AAGAATATAA AAGACAATTT 6360
GCAACAAATC TCAGGTCGGA TTGATATTAT TCACAAGAAG AAGACAGCAG CCTTGCAAAG 6420
TGCCACCTCC ATGGAAAAGG TGAAAGTACA GGAAGCCGTG GCACAGATGG ATTTCCAGGG 6480
GGAAAAACTT CATAGAATGT ACAAGGAACG ACAAGGGCGA TTCGACAGAT CAGTTGAAAA 6540
ATGGCGACAC TTTCATTATG ATATGAAGGT ATTTAATCAA TGGCTGAATG AAGTTGAACA 6600
GTTTTTCAAA AAGACACAAA ATCCTGAAAA CTGGGAACAT GCTAAATACA AATGGTATCT 6660
TAAGGAACTC CAGGATGGCA TTGGGCAGCG TCAAGCTGTT GTCAGAACAC TGAATGCAAC 6720
TGGGGAAGAA ATAATTCAAC AGTCTTCAAA AACAGATGTC AATATTCTAC AAGAAAAATT 6780
AGGAAGCTTG AGTCTGCGGT GGCACGACAT CTGCAAAGAG CTGGCAGAAA GGAGAAAGAG 6840
GATTGAAGAA CAAAAGAATG TCTTGTCAGA ATTTCAAAGA GATTTAAATG AATTTGTTTT 6900
GTGGCTGGAA GAAGCAGATA ACATTGCTAT TACTCCACTT GGAGATGAGC AGCAGCTAAA 6960
AGAACAACTT GAACAAGTCA AGTTACTGGC AGAAGAGTTG CCCCTGCGCC AGGGAATTCT 7020
AAAACAATTA AATGAAACAG GAGGAGCAGT ACTTGTAAGT GCTCCCATAA GGCCAGAAGA 7080
GCAAGATAAA CTTGAAAAGA AGCTCAAACA GACAAATCTC CAGTGGATAA AGGTCTCCAG 7140
AGCTTTACCT GAGAAACAAG GAGAGCTTGA GGTTCACTTA AAAGATTTTA GGCAGCTTGA 7260
AGAGCAGCTG GATCACCTGC TTCTGTGGCT CTCTCCTATT AGAAACCAGT TGGAAATTTA

TAACCAACCA AGTCAGGCAG GACCGTTTGA CATAAAGGAG ATTGAAGTAA CAGTTCACGG 7320
TAAACAAGCG GATGTGGAAA GGCTTTTGTC GAAAGGGCAG CATTTGTATA AGGAAAAACC 7380
AAGCACTCAG CCAGTGAAGA GGAAGTTAGA AGATCTGAGG TCTGAGTGGG AGGCTGTAAA 7440
CCATTTACTT CGGGAGCTGA GGACAAAGCA GCCTGACCGT GCCCCTGGAC TGAGCACTAC 7500
TGGAGCCTCT GCCAGTCAGA CTGTTACTCT AGTGACACAA TCTGTGGTTA CTAAGGAAAC 7560
TGTCATCTCC AAACTAGAAA TGCCATCTTC TTTGCTGTTG GAGGTACCTG CACTGGCAGA 7620
CTTCAACCGA GCTTGGACAG AACTTACAGA CTGGCTGTCT CTGCTTGATC GAGTTATAAA 7680
ATCACAGAGA GTGATGGTGG GTGATCTGGA AGACATCAAT GAAATGATCA TCAAACAGAA 7740
GGCAACACTG CAAGATTTGG AACAGAGACG CCCCCAATTG GAAGAACTCA TTACTGCTGC 7800
CCAGAATTTG AAAAACAAAA CCAGCAATCA AGAAGCTAGA ACAATCATTA CTGATCGAAT 7860
TGAAAGAATT CAGATTCAGT GGGATGAGGT TCAAGAACAG CTGCAGAACA GGAGACAACA 7920
GTTGAATGAA ATGTTAAAGG ATTCAACACA ATGGCTGGAA GCTAAGGAAG AAGCCGAACA 7980
GGTCATAGGA CAGGTCAGAG GCAAGCTTGA CTCATGGAAA GAAGGTCCTC ACACAGTAGA 8040
TGCAATCCAA AAGAAGATCA CAGAAACCAA GCAGTTGGCC AAAGACCTCC GTCAACGGCA 8100
GATAAGTGTA GACGTGGCAA ATGATTTGGC ACTGAAACTT CTTCGGGACT ATTCTGCTGA 8160
TGATACCAGA AAAGTACACA TGATAACAGA GAATATCAAT ACTTCTTGGG GAAACATTCA 8220
TAAAAGAGTA AGTGAGCAAG AGGCTGCTTT GGAAGAAACT CATAGATTAC TGCAGCAGTT 8280
CCCTCTGGAC CTGGAGAAGT TTCTTTCCTG GATTACGGAA GCAGAAACAA CTGCCAATGT 8340
CCTACAGGAC GCTTCCCGTA AGGAGAAGCT CCTAGAAGAC TCCAGGGGAG TCAGAGAGCT 8400
GATGAAACCA TGGCAAGATC TCCAAGGAGA AATTGAAACT CACACAGATA TCTATCACAA 8460
TCTTGATGAA AATGGCCAAA AAATCCTGAG ATCCCTGGAA GGTTCGGATG AAGCACCCCT 8520
GTTACAAAGA CGTTTGGATA ACATGAATTT CAAGTGGAGT GAACTTCAGA AAAAGTCTCT 8580
CAACATTAGG TCCCATTTGG AAGCAAGTTC TGACCAGTGG AAGCGTTTGC ATCTTTCTCT 8640
TCAGGAACTT CTTGTTTGGC TACAGCTGAA AGATGATGAA CTGAGCCGTC AGGCACCCAT 8700
CGGTGGTGAT TTCCCAGCAG TTCAGAAGCA GAATGATATA CATAGGGCCT TCAAGAGGGA 8760
ATTGAAAACT AAAGAACCTG TAATCATGAG TACTCTGGAG ACTGTGAGAA TATTTCTGAC 8820
AGAGCAGCCT TTGGAAGGAC TAGAGAAACT CTACCAGGAG CCCAGAGAAC TGCCTCCTGA 8880
AGAAAGAGCT CAGAATGTCA CTCGGCTCCT ACGAAAGCAG GCTGAAGAGG TCAACGCTGA 8940
ATGGGACAAA TTGAACCTGC GCTCAGCTGA TTGGCAGAGA AAAATAGATG AAGCTCTTGA 9000
AAGACTCCAG GAACTTCAGG AAGCTGCCGA TGAACTGGAC CTCAAGTTGC GCCAAGCTGA 9060
GGTGATCAAG GGATCCTGGC AGCCAGTGGG GGATCTCCTC ATTGACTCTC TGCAAGATCA 9120
CCTTGAAAAA GTCAAGGCAC TTCGGGGAGA AATTGCACCT CTTAAAGAGA ATGTCAATCG 9180
TGTCAATGAC CTTGCACATC AGCTGACCAC ACTGGGCATT CAGCTCTCAC CTTATAACCT 9240
CAGCACTTTG GAAGATCTGA ATACCAGATG GAGGCTTCTA CAGGTGGCTG TGGAGGACCG 9300
TGTCAGACAG CTGCATGAAG CCCACAGGGA CTTTGGTCCT GCATCCCAGC ACTTCCTTTC 9360
CACTTCAGTT CAGGGTCCCT GGGAGAGAGC CATCTCACCA AACAAAGTGC CCTACTATAT 9420
CAACCACGAG ACCCAAACCA CTTGTTGGGA CCACCCCAAA ATGACAGAGC TCTACCAGTC 9480
TTTAGCTGAC CTGAATAATG TCAGGTTCTC CGCGTATAGG ACTGCCATGA AGCTCAGAAG 9540
GCTCCAGAAG GCCCTTTGCT TGGATCTCTT GAGCCTGTCA GCTGCATGTG ATGCCCTGGA 9600
CCAGCACAAC CTCAAGCAAA ATGACCAGCC CATGGATATC CTGCAGATAA TTAACTGTTT 9660
GACTACAATT TATGATCGTC TGGAGCAAGA GCACAACAAT CTGGTCAATG TCCCTCTCTG 9720
TGTGGATATG TGTCTCAACT GGCTTCTCAA TGTTTATGAT ACGGGACGAA CAGGGAGGAT 9780
CCGTGTCCTG TCTTTTAAAA CTGGCATCAT TTCTCTGTGT AAAGCACACT TGGAAGACAA 9840
GTACAGATAC CTTTTCAAGC AAGTGGCAAG TTCAACTGGC TTTTGTGACC AGCGTAGGCT 9900
GGGTCTTCTT CTGCATGATT CTATTCAAAT CCCAAGACAG TTGGGTGAAG TTGCTTCCTT 9960
TGGGGGCAGT AACATTGAGC CGAGTGTCAG GAGCTGCTTC CAATTTGCCA ATAATAAACC 10020
TGAGATTGAA GCTGCTCTCT TCCTTGACTG GATGCGCCTG GAACCCCAGT CTATGGTGTG 10080

GCTGCCCGTC TTGCACAGAG TGGCTGCTGC TGAAACTGCC AAGCATCAAG CCAAGTGTAA 10140
CATCTGTAAG GAGTGTCCAA TCATTGGATT CAGGTACAGA AGCCTAAAGC ATTTTAATTA 10200
TGACATCTGC CAAAGTTGCT TTTTTTCTGG CCGAGTTGCA AAGGGCCATA AAATGCACTA 10260
CCCCATGGTA GAGTATTGCA CTCCGACTAC ATCCGGAGAA GATGTTCGCG ACTTCGCCAA 10320
GGTACTAAAA AACAAATTTC GAACCAAAAG GTATTTTGCG AAGCATCCCC GAATGGGCTA 10380
CCTGCCAGTG CAGACTGTGT TAGAGGGGGA CAACATGGAA ACTCCCGTTA CTCTGATCAA 10440
CTTCTGGCCA GTAGATTCTG CGCCTGCCTC GTCCCCCCAG CTTTCACACG ATGATACTCA 10500
TTCACGCATT GAACATTATG CTAGCAGGCT AGCAGAAATG GAAAACAGCA ATGGATCTTA 10560
TCTAAATGAT AGCATCTCTC CTAATGAGAG CATAGATGAT GAACATTTGT TAATCCAGCA 10620
TTACTGCCAA AGTTTGAACC AGGACTCCCC CCTGAGCCAG CCTCGTAGTC CTGCCCAGAT 10680
CTTGATTTCC TTAGAGAGTG AGGAAAGAGG GGAGCTAGAG AGAATCCTAG CAGATCTTGA 10740
GGAAGAAAAC AGGAATCTGC AAGCAGAATA TGATCGCCTG AAGCAGCAGC ATGAGCATAA 10800
AGGCCTGTCT CCACTGCCAT CTCCTCCTGA GATGATGCCC ACCTCTCCTC AGAGTCCCAG 10860
GGATGCTGAG CTCATTGCTG AGGCTAAGCT ACTGCGCCAA CACAAAGGAC GCCTGGAAGC 10920
CAGGATGCAA ATCCTGGAAG ACCACAATAA ACAGCTGGAG TCTCAGTTAC ATAGACTGAG 10980
ACAGCTCCTG GAGCAGCCCC AGGCTGAAGC TAAGGTGAAT GGCACCACGG TGTCCTCTCC 11040
TTCCACCTCT CTGCAGAGGT CAGATAGCAG TCAGCCTATG CTGCTCCGAG TGGTTGGCAG 11100
TCAAACTTCA GAATCTATGG GTGAGGAAGA TCTTCTGAGT CCTCCCCAGG ACACAAGCAC 11160
AGGGTTAGAA GAAGTGATGG AGCAACTCAA CAACTCCTTC CCTAGTTCAA GAGGAAGAAA 11220
TGCCCCCGGA AAGCCAATGA GAGAGGACAC AATGTAGGAA GCCTTTTCCA CATGGCAGAT 11280
GATTTGGGCA GAGCGATGGA GTCCTTAGTT TCAGTCATGA CAGATGAAGA AGGAGCAGAA 11340
TAAATGTTTT ACAACTCCTG ATTCCCGCAT GGTTTTTATA ATATTCGTAC AACAAAGAGG 11400
ATTAGACAGT AAGAGTTTAC AAGAAATAAA ATCTATATTT TTGTGAAGGG TAGTGGTACT 11460
ATACTGTAGA TTTCAGTAGT TTCTAAGTCT GTTATTGTTT TGTTAACAAT GGCAGGTTTT 11520
ACACGTCTAT GCAATTGTAC AAAAAAGTTA AAAGAAAACA TGTAAAATCT TGATAGCTAA 11580
ATAACTTGCC ATTTCTTTAT ATGGAACGCA TTTTGGGTTG TTTAAAAATT TATAACAGTT 11640
ATAAAGAGAG ATTGTAAACT AAAGTGTGCT TTATAAAAAA AGTTGTTTAT AAAAACCCCT 11700
AAACAAACAC ACACGCACAC ACACACACAC ACACACACAC ACACACACAC GCACACATAC 11760
ATGCACGAAC CCACCACACA CACACACACA CACACACACA CTGAGGCAGC ACATTGTTTT 11820
GCATTACTTT AGCGTGGTAT TCATATGGAA TTCATGACGT TTTTTTATTT TCTTGCATAC 11880
GAACCCCACC AAATGACTGC TTCATATTGC TCTTTTGAGA ATTGTTGACT GAGTGGGGCT 11940
GGCTATGGGC TTTCATTTTA TACATCTATA TGTCTACAAG TATATAAATA CTATAGGTAT 12000
ATAGATAAAT AGATATGAAG TTACTTCTTC AAATGTTCTT GCCACTTCCT AATGGAAATT 12060
GCTTCTAGTC ATCTGGGCTT ATCTGCTTGG GCAAGAGTGA ATTTTCCCTG GAGCCCAAAG 12120
CCAGGAGACT ACCGCCACAC TAAAATATTG TCTAGGGCTC CAGATGTTTC TAGTTTTAAA 12180
CTTTCCACTG AGAGCTAGAG GATTCATTTT TTTCAAGGAA CATGCGAATG AATACACAGG 12240
ACTTACTATC ATAGTAATTT GTTGGCTGAT ATATTCAACT TCCTACTGTT GGGTTATATT 12300
TAATGATGTT TCTGCAATAG AACATCAGAT GACATTTTTA ACTCCCAGAC AGTAGGAGGA 12360
AGATGGTAGG AGCTAAAGGT TGCGGCTCCT CAGTCAATTT ATATGAGGGG AGCAACAACT 12420
CTGTAAAAGA ATGGATGAAT ATTTACAACT ATACATATAA ACATCTCTAT AATTACAACT 12480
AAATTGTTCT GCCCTCTTCA TAAACTCAAC CTGAAGTGGG TGGTTTTGTT GTTGTTGTTG 12540
TTGTTGTTGT TGATGATGAT GATGAATTTT AGATTTTAGA TTTTTTGGGT TTTTTTTTCT 12600
TCATTGTGAT GATTTTTTTT TTTAATGCTG CAAGACTTAG GATTACTGTT AAGAAAGTAA 12660
CCCAATCACA TTGTGACCCT GGTGAATATC AGTCCAGAAG CCCATGAACT GCATTTGTCT 12720
CCTTTGCATT GGTTTCCCTG CAAGTAACTC CACACAGGAT TGTGGGTGAG AAGGCACAGT 12780
GGTTGGAAAG TTTTGAGAGC AAAAGCGTCT CCAAACTCTC TGGTCTAGTT GACGGGCTGA 12840
AATGTCTAAA CAAATGCAAG TCATTGAACC AGGAGAAAAA GTGCAACAGA AAGCTAAGGA 12900

CTGCTAGGAA GAGCTTTACT CCTCTCATGC CAGTTTCTTC TTCTTAGCAT TTAAAGAGCA 12960
TTCTCTCAAT AGAAATCACT GTCCTATCAT TTTGCAAATC TGTTACCTCT AACGTCAAGT 13020
GTAATTAACT TCTAGCGAGT GGGTTTTGTC CATTATTAAT TGTAATTAAC ATCAAACACA 13080
GCTTCTCATG CTATTTCTAC CTCACTTTGG TTTTGGGGTG TTTCTAGTAA TTGTGCACAC 13140
CTAATTTCAC AACTTCACCA CTTGTCTGTT GTGTGGACAC CAGTTTCCTT TTTTCATTTA 13200
TAATTTCCAA AAGAAAACCC AAAGCTCTAA GATAACAAAT TGAAATTTGG TTCTGGTCTT 13260
GCTTTTCTCT CTCTCTCTCC TTTATGTGGC ACTGGGCATT TTCTTTATCC AAGGATTTGT 13320
TTTCACCAAG ATTTAAAACA AGGGGTTCCT TTCCTACTAA GAAGTTTTAA GTTTCATTCT 13380
AAAATCCAAG GTAGATAGAG TGCATAGTTT TGTTTTAATC TTTTCGTTTT ATCTTTTAGA 13440
TATTAGTTCT GGAGTGAATC TATCAAAATA TTTGAATAAA AACTGAGAGC TTTATTGCTG 13500
ATTTTAAGCA TAATTTGGAC ATCATTTCAT GTTCTTTATA ACCATCAAGT ATTAAAGTGT 13560
AAATCATAAT CAGTGTAACT GAAGCATAAT CATCACATGG CATGTATCAT CATTGTCTCC 13620
AGGTACTGGA CTCTTACTTG AGTATCATAA TAGATTGTGT TTTAACACCA ACACTGTAAC 13680
ATTTACTAAT TATTTTTTTA AACTTCAGTT TTACTGCATT TTCACAACAT ATCAGATTTC 13740
ACCAAATATA TGCCTTACTA TTGTATTATA TTACTGCTTT ACTGTGTATC TCAATAAAGC 13800
ACGCAGTTAT GTTAC 13815

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

ATACCGAGGC TGCAGTGTAC A 21

(2) INFORMATION FOR SEQ ID NO:19:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 50 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GCAACCGAAC ATATCGCCAC GGCAGCCACC AACGATGCTG ATTGCCGTTC 50

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 57 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

GGTGGCTGCC GTGGCGATAT GTTCGGTTGC TAACTTTACC CTTCATCACT AAAGGCC 57

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

AAACGTACAG CGCCATGTTT ACCAG 25

CLAIMS

1. In a method for gene therapy, the improvement which comprises using a bacteriophage containing genetic material to introduce a therapeutic polynucleotide into a target cell of a mammalian recipient, wherein the mammalian recipient is diagnosed as having a condition that is treatable by administration to the recipient of the therapeutic polynucleotide or a product thereof, the method comprising the steps of:

(1) contacting the bacteriophage with the target cell under conditions to permit selective binding of a ligand on the surface of the bacteriophage to a receptor on the surface of the target cell and to allow the bacteriophage to enter the target cell; and

(2) allowing the target cell to live under conditions such that the therapeutic polynucleotide is transcribed therein,

wherein the bacteriophage is incapable of injecting said genetic material into the target cell of the mammalian recipient.

2. The method of claim 1, wherein the bacteriophage propagates in a prokaryotic cell.

3. The method of claim 1, wherein the bacteriophage includes exogenous genetic material that is transcribed and translated in the target cell.

4. The method of claim 1, wherein the bacteriophage is selected from the group consisting of a lambda phage, a P1 phage, a P22 phage, a T1 phage, a T2 phage, a T3 phage, a T4 phage, a T5 phage, a T6 phage, a T7 phage, a P2 phage, a P4 phage, an Mu phage, a PM2 phage, an N4 phage, an SPO1 phage, a PBS1 phage, and a PBS2 phage.

5. The method of claim 1, wherein the therapeutic polynucleotide is operably coupled to a promoter.

6. The method of claim 1, wherein contacting the bacteriophage with the target cell is performed in vitro.

7. The method of claim 1, wherein the bacteriophage contains exogenous genetic material that contains a cell-specific promoter that effects transcription and translation of the therapeutic polynucleotide in the target cell.

8. The method of claim 1, wherein the bacteriophage contains one or more modified tail proteins, wherein the modified tail proteins cannot facilitate injection of the genetic material into the target cell of the mammalian recipient.

9. A method for introducing an exogenous polynucleotide into a mammalian cell comprising:

(1) contacting a bacteriophage containing genetic material with the target cell under conditions to permit selective binding of a ligand on the surface of the bacteriophage to a receptor on the surface of the target cell and to allow the bacteriophage to enter the target cell; and

(2) allowing the target cell to live under conditions such that the therapeutic polynucleotide is transcribed therein,

wherein the bacteriophage is incapable of injecting the genetic material into the target cell.

10. The method of claim 9, wherein the bacteriophage includes exogenous genetic material that is transcribed and translated in the target cell.

11. The method of claim 9, wherein the bacteriophage is derived from a bacteriophage selected from the group consisting of a lambda phage, a P1 phage, a P22 phage, a T1 phage, a T2 phage, a T3 phage, a T4 phage, a T5 phage, a T6 phage, a T7 phage, a P2 phage, a P4 phage, an Mu phage, a PM2 phage, an N4 phage, an SPO1 phage, a PBS1 phage, and a PBS2 phage.

12. The method of claim 9, wherein the exogenous polynucleotide comprises a therapeutic polynucleotide.

13. The method of claim 9, wherein the bacteriophage contains one or more modified tail proteins, wherein the modified tail proteins cannot facilitate injection of the genetic material into the target cell of the mammalian recipient.

14. The method of claim 9, wherein contacting the bacteriophage with the target cell is performed in vitro.

15. The method of claim 9, wherein the bacteriophage has a genome that contains a cell-specific promoter that effects transcription and translation of the therapeutic polynucleotide in the target cell.

16. A bacteriophage comprising:

(a) a bacteriophage genome containing an exogenous polynucleotide that can be transcribed in a mammalian cell; and

(b) a ligand contained on the surface of the bacteriophage, wherein the ligand

selectively binds to a receptor expressed on the surface of a mammalian cell,

wherein the bacteriophage is incapable of injecting the bacteriophage genome into the mammalian cell.

17. The composition of claim 16, wherein the composition is contained in an implant that is suitable for implantation into a mammalian recipient

18. A bacteriophage comprising:

a bacteriophage containing a bacteriophage genome that can be transcribed in a mammalian cell, wherein the bacteriophage contains avidin on its surface and wherein the bacteriophage is incapable of injecting the bacteriophage genome into the mammalian cell.

19. The bacteriophage of claim 18, wherein the bacteriophage is contained in a kit with instructions for attaching a biotinylated ligand to the bacteriophage to form a ligand-labeled bacteriophage.

20. A kit comprising a container including:

(1) a first container containing a bacteriophage having a surface and a genome, wherein the bacteriophage is incapable of injecting the bacteriophage genome into a mammalian cell;

(2) a second container containing an agent for attaching a ligand to the surface of a bacteriophage, wherein the ligand binds to a receptor on the surface of the mammalian cell; and

(3) instructions for attaching the agent to the surface of the bacteriophage.

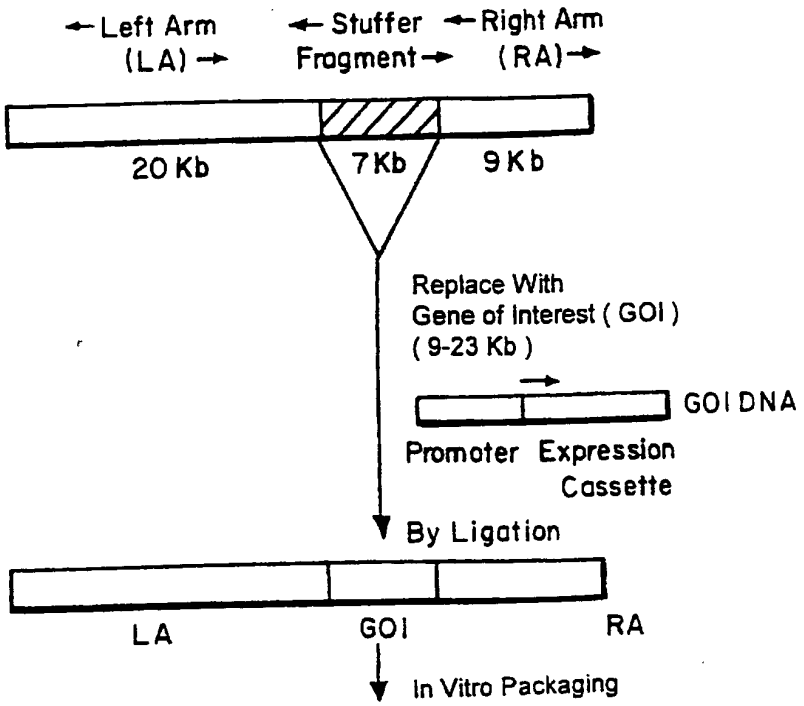

Left Arm (LA)
Stuffer Fragment
Right Arm (RA)
20 Kb
7 Kb
9 Kb
Replace With Gene of Interest (GOI) (9-23 Kb)
GOI DNA
Promoter
Expression Cassette
By Ligation
LA
GOI
RA
In Vitro Packaging

FIG. 1A

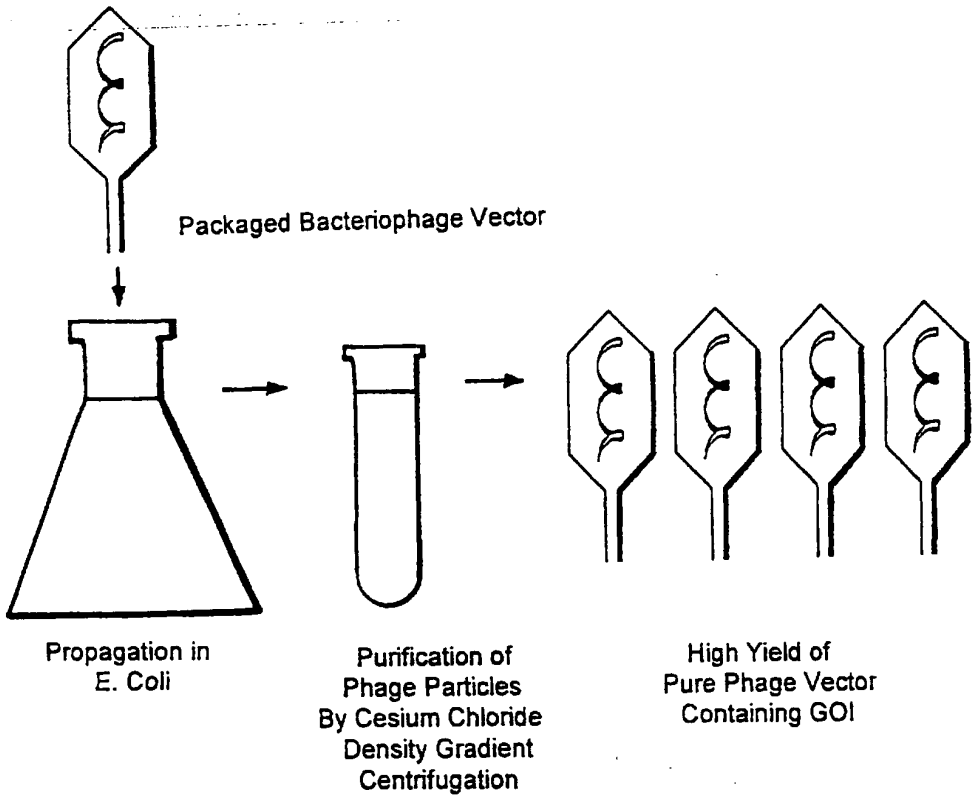
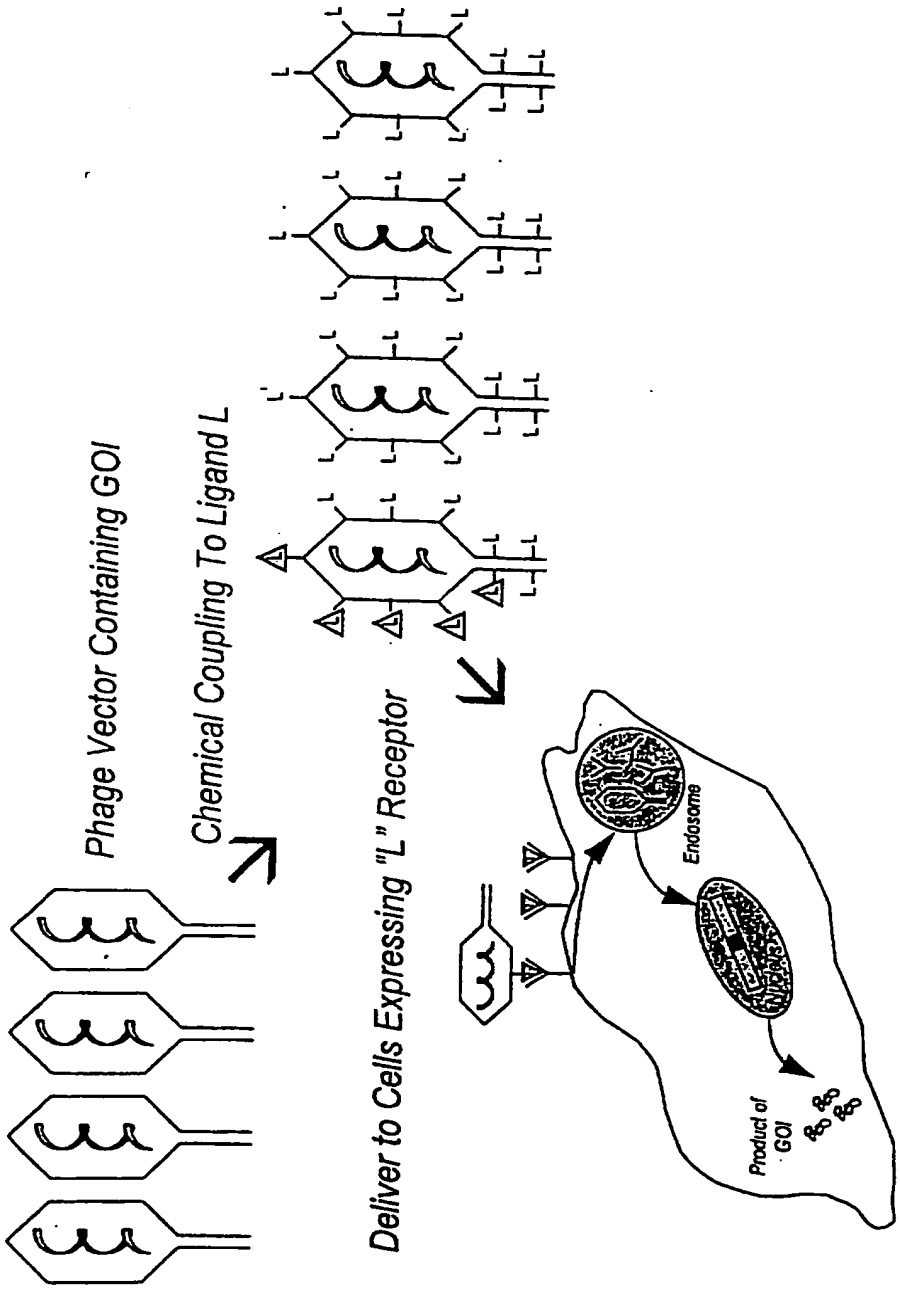

Packaged Bacteriophage Vector
Propagation in E. Coli
Purification of Phage Particles By Cesium Chloride Density Gradient Centrifugation
High Yield of Pure Phage Vector Containing GOI

FIG. 1B

Figure 1C


Phage Vector Containing GOI
Chemical Coupling To Ligand L
Deliver to Cells Expressing "L" Receptor
Endosome
Product of GOI

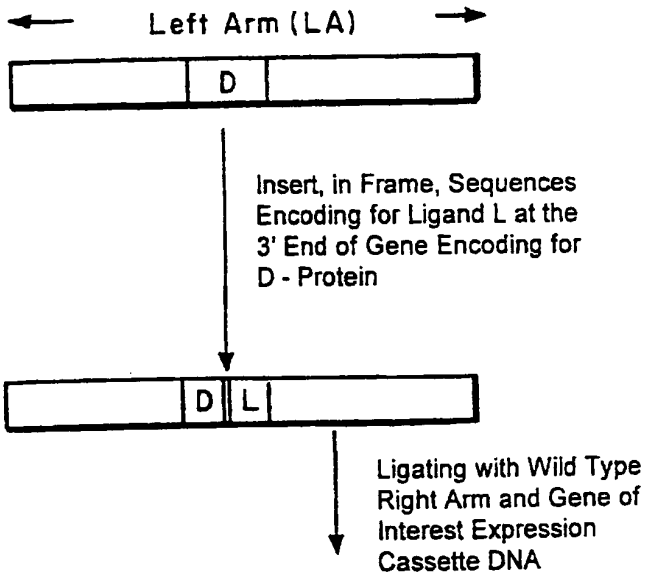

Left Arm (LA)
D
Insert, in Frame, Sequences Encoding for Ligand L at the 3' End of Gene Encoding for D - Protein
D
L
Ligating with Wild Type Right Arm and Gene of Interest Expression Cassette DNA

FIG. 2

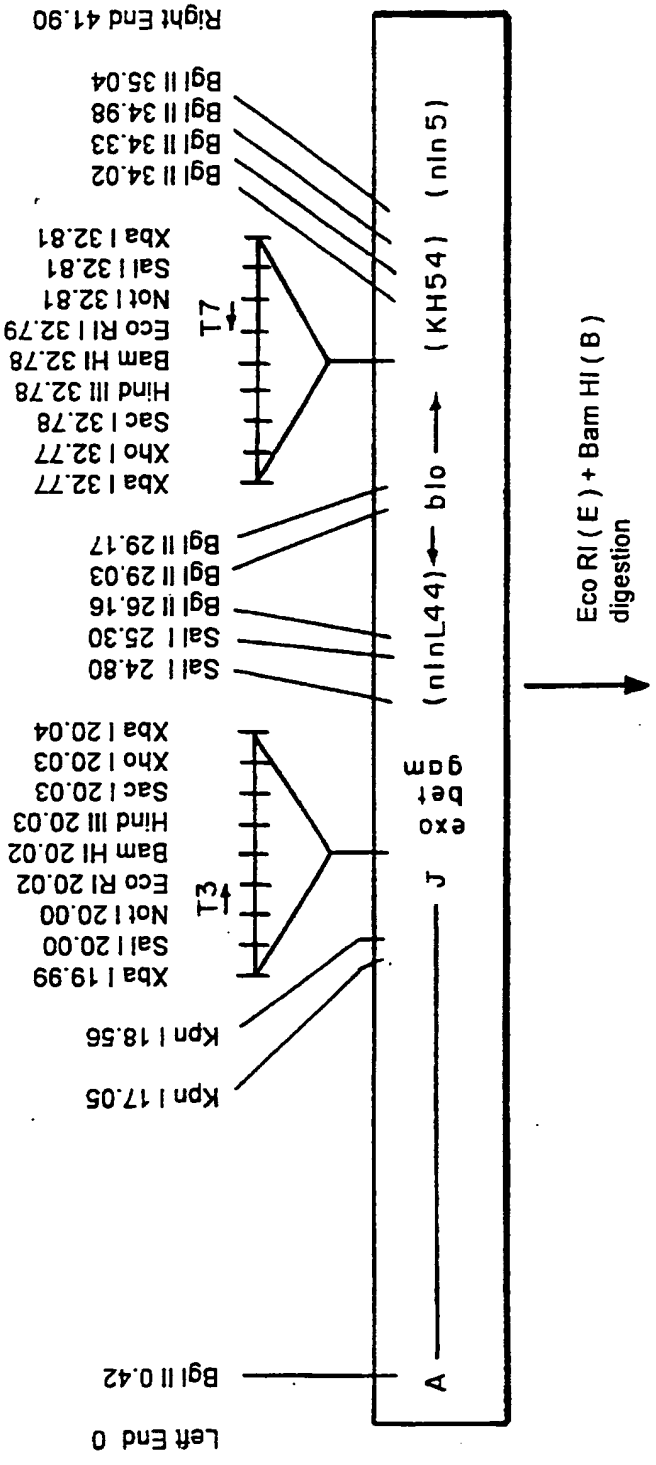

Left End 0
Bgl II 0.42
A
J
exo
bet
gam
(nInL44)
blo
(KH54)
(nIn5)
Kpn I 17.05
Kpn I 18.56
Xba I 19.99
Sal I 20.00
Not I 20.00
Eco RI 20.02
Bam HI 20.02
Hind III 20.03
Sac I 20.03
Xho I 20.03
Xba I 20.04
T3
Sal I 24.80
Sal I 25.30
Bgl II 26.16
Bgl II 29.03
Bgl II 29.17
Xba I 32.77
Xho I 32.77
Sac I 32.78
Hind III 32.78
Bam HI 32.78
Eco RI I 32.79
Not I 32.81
Sal I 32.81
Xba I 32.81
T7
Bgl II 34.02
Bgl II 34.33
Bgl II 34.98
Bgl II 35.04
Right End 41.90
Eco RI (E) + Bam HI (B)
digestion

FIG. 3A

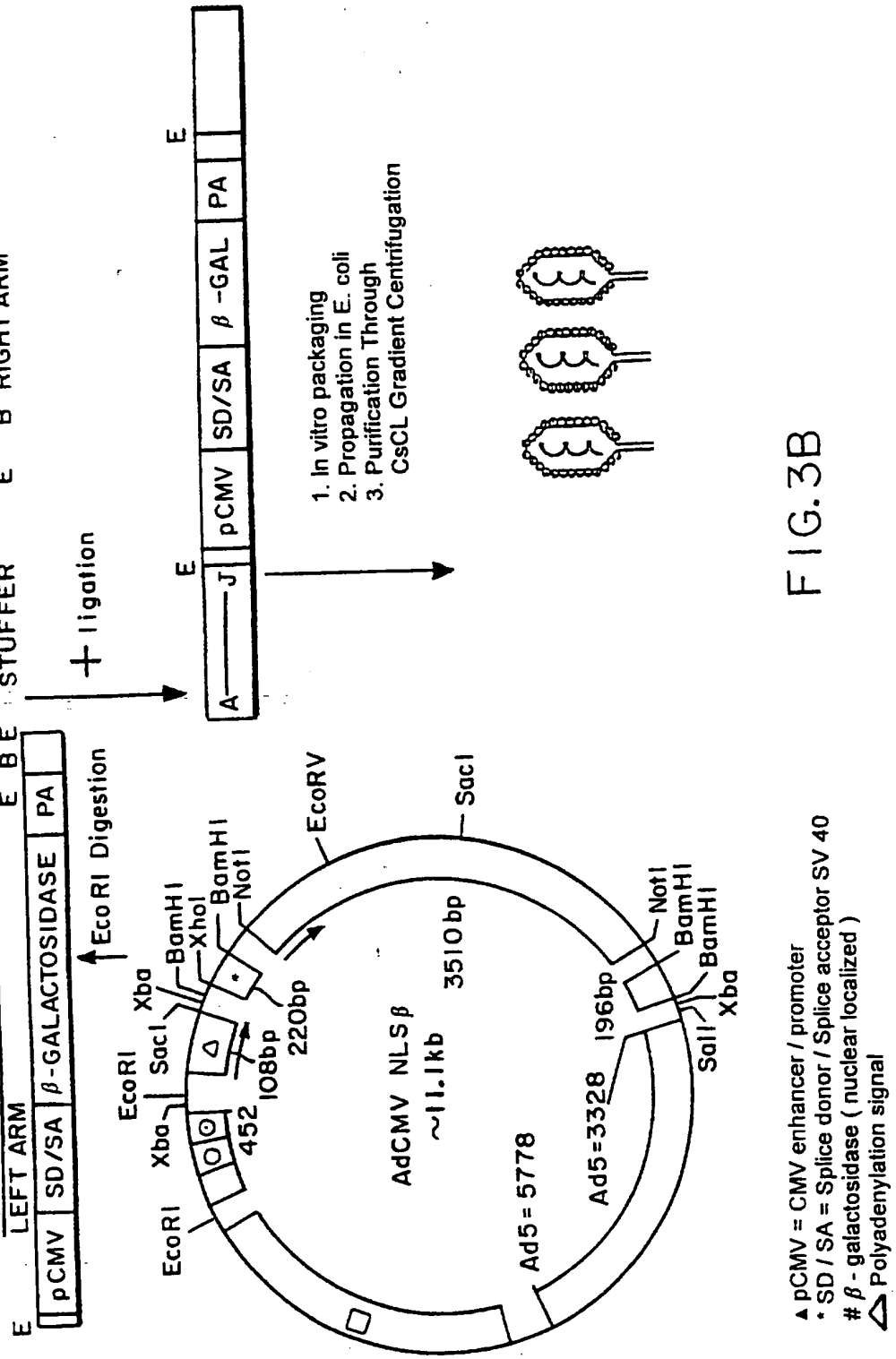

E B
E B
A J E B E STUFFER E B RIGHT ARM
LEFT ARM
E pCMV SD/SA β-GALACTOSIDASE PA
EcoRI Digestion
+ ligation
E A J pCMV SD/SA β-GAL PA E
1. In vitro packaging
2. Propagation in E. coli
3. Purification Through
CsCL Gradient Centrifugation
EcoRV
SacI
BamHI
XhoI
NotI
Xba
EcoRI
SacI
Xba
EcoRI
452
108bp
220bp
AdCMV NLSβ
~11.1kb
3510bp
Ad5=5778
Ad5=3328
196bp
NotI
BamHI
BamHI
SalI
Xba
▲ pCMV = CMV enhancer / promoter
* SD / SA = Splice donor / Splice acceptor SV 40
# β - galactosidase (nuclear localized)
△ Polyadenylation signal

FIG. 3B

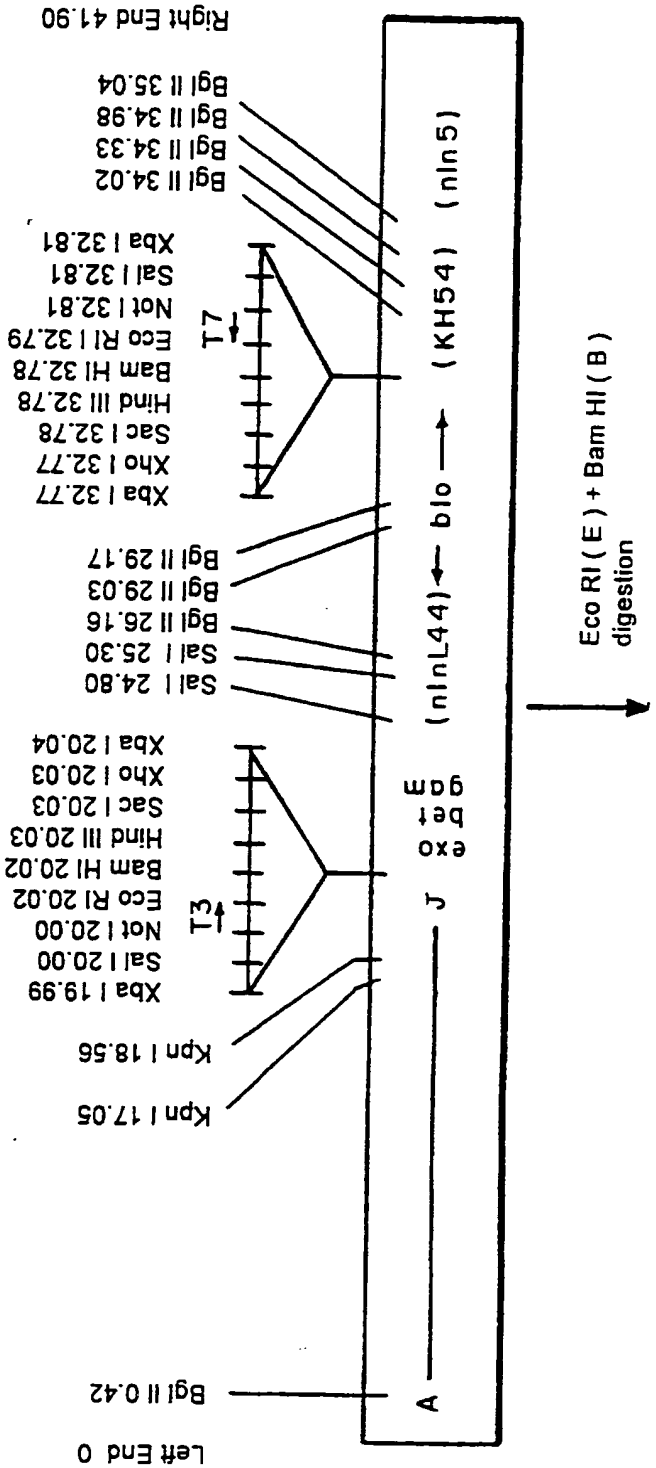

Left End 0
Bgl II 0.42
A
J
exo bet gam
(nInL44)
blo
(KH54)
(nIn5)
Kpn I 17.05
Kpn I 18.56
Xba I 19.99
Sal I 20.00
Not I 20.00
Eco RI 20.02
Bam HI 20.02
Hind III 20.03
Sac I 20.03
Xho I 20.03
Xba I 20.04
T3
Sal I 24.80
Sal I 25.30
Bgl II 26.16
Bgl II 29.03
Bgl II 29.17
Xba I 32.77
Xho I 32.77
Sac I 32.78
Hind III 32.78
Bam HI 32.78
Eco RI I 32.79
Not I 32.81
Sal I 32.81
Xba I 32.81
T7
Bgl II 34.02
Bgl II 34.33
Bgl II 34.98
Bgl II 35.04
Right End 41.90
Eco RI (E) + Bam HI (B)
digestion

FIG. 4A

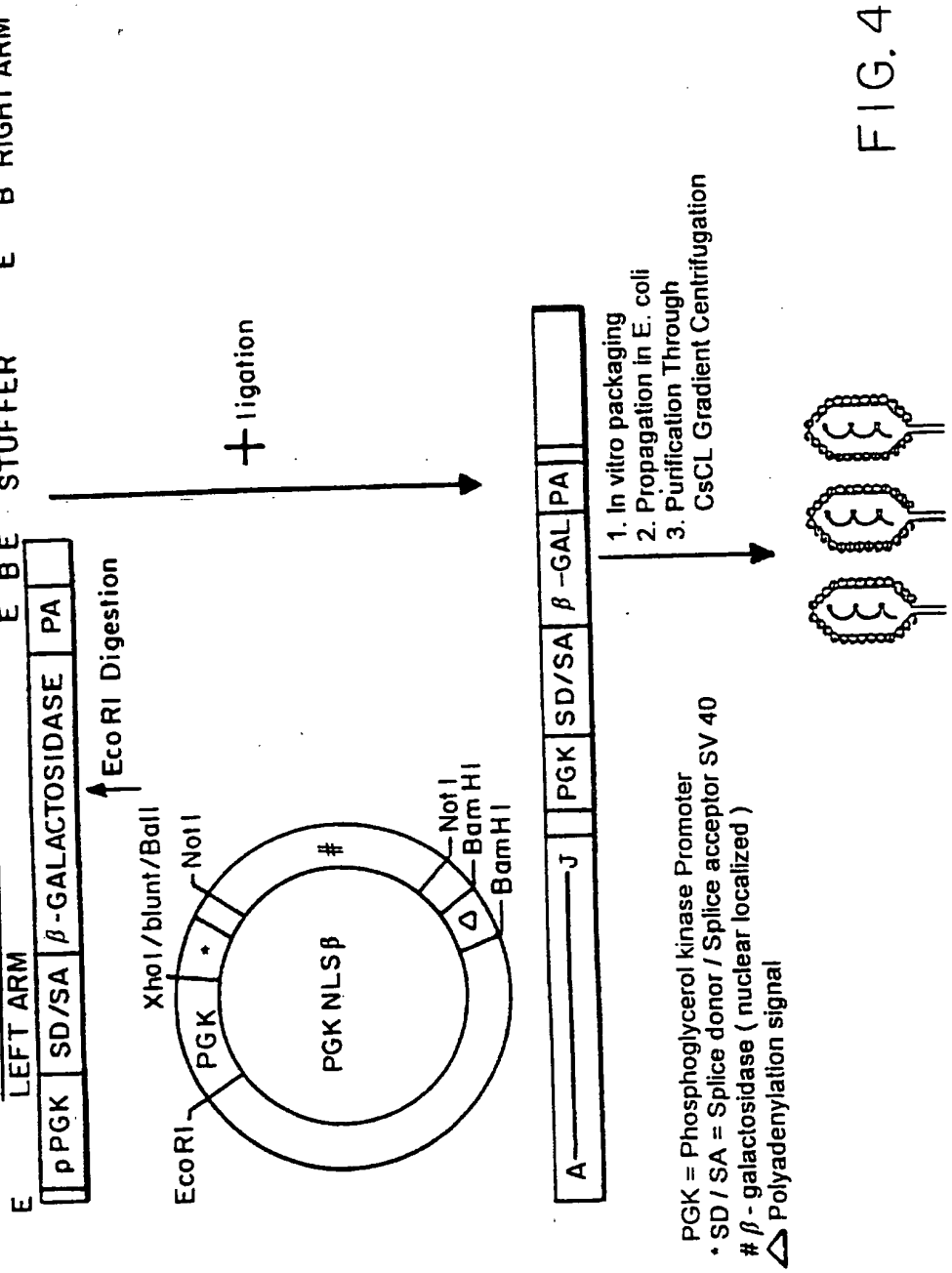

E
pPGK
SD/SA
β-GALACTOSIDASE
PA
LEFT ARM
A
J
E
B
E
B
E
STUFFER
E
B
RIGHT ARM
Eco RI Digestion
+ ligation
EcoRI
PGK
*
Xho I / blunt / Ball
Not I
PGK NLSβ
#
Not I
Bam HI
Δ
Bam HI
A
J
PGK
SD/SA
β-GAL
PA
1. In vitro packaging
2. Propagation in E. coli
3. Purification Through
CsCL Gradient Centrifugation
PGK = Phosphoglycerol kinase Promoter
* SD / SA = Splice donor / Splice acceptor SV 40
# β - galactosidase (nuclear localized)
Δ Polyadenylation signal

FIG. 4B

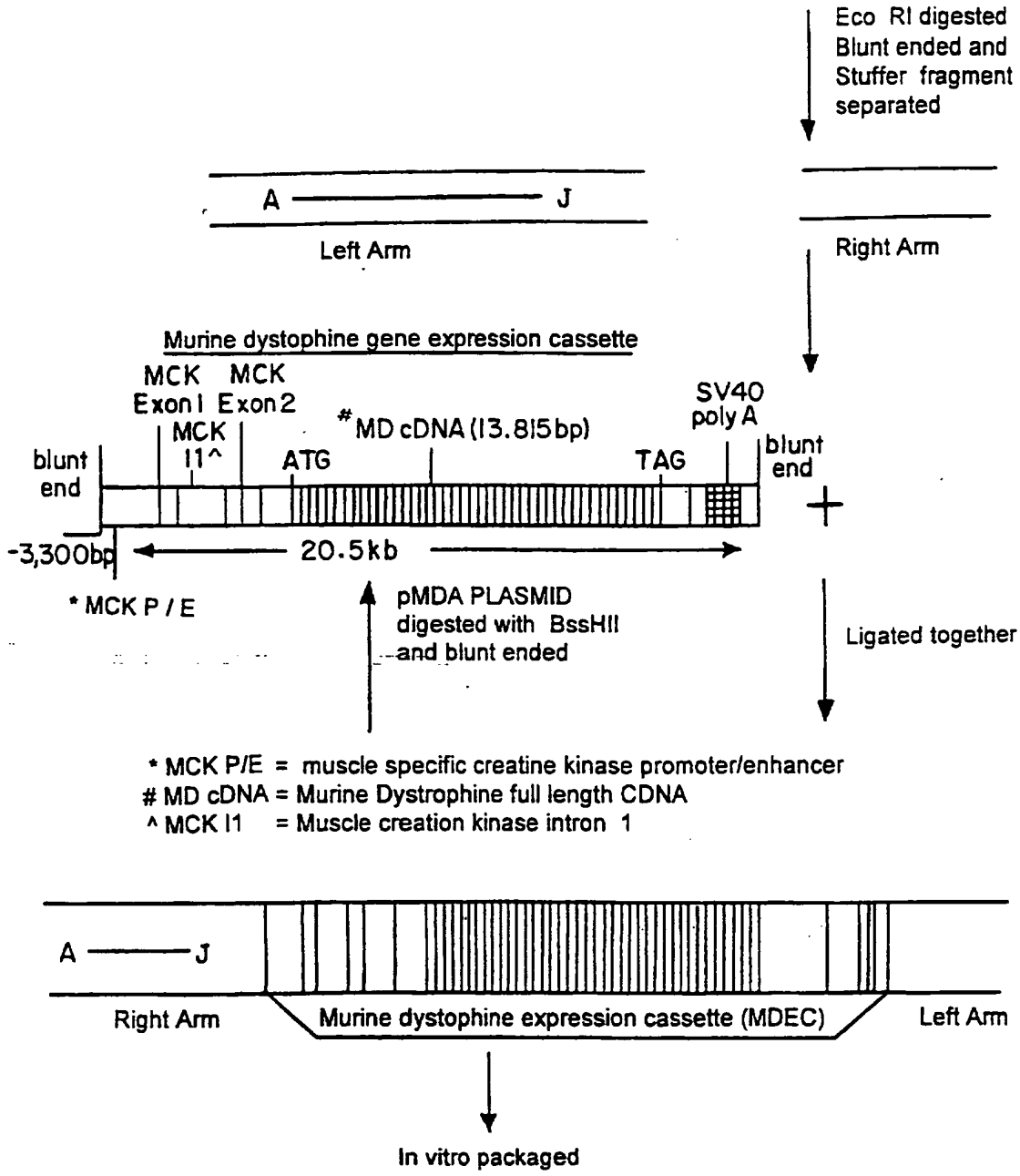
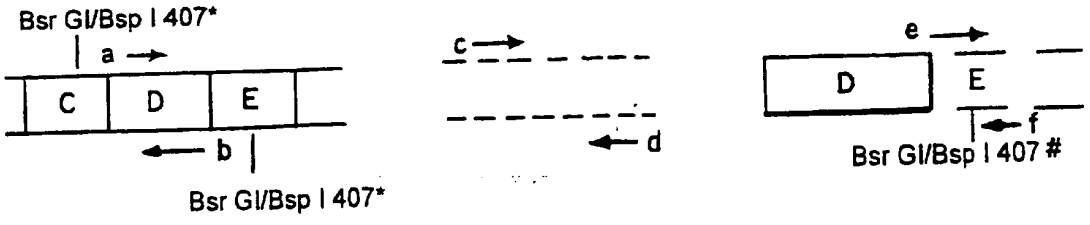

Eco RI digested
Blunt ended and
Stuffer fragment
separated
A
J
Left Arm
Right Arm
Murine dystophine gene expression cassette
MCK Exon1
MCK Exon2
MCK I1^
MD cDNA (13.815bp)
ATG
TAG
SV40 poly A
blunt end
blunt end
-3,300bp
20.5kb
* MCK P / E
pMDA PLASMID
digested with BssHII
and blunt ended
Ligated together
* MCK P/E = muscle specific creatine kinase promoter/enhancer
# MD cDNA = Murine Dystrophine full length CDNA
^ MCK I1 = Muscle creation kinase intron 1
A
J
Right Arm
Murine dystophine expression cassette (MDEC)
Left Arm
In vitro packaged

FIG. 5



Wt λ genomic DNA template Heregulin cDNA template Wt λ genomic DNA template

* the location of this site is position 5220
\# the location of this site is position 6140

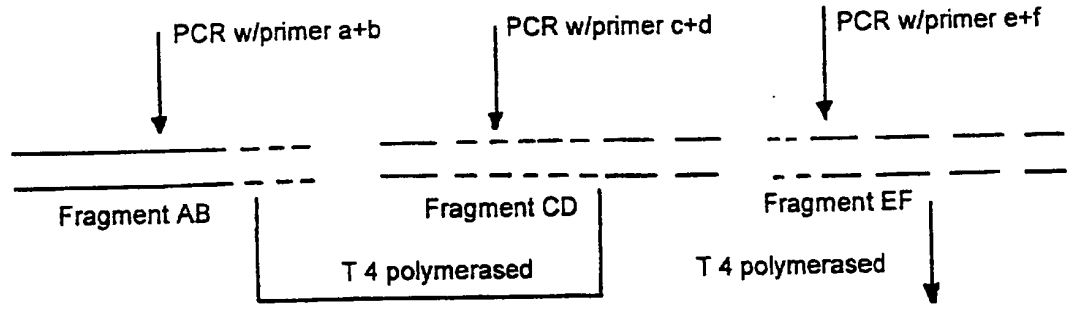


FIG. 6A

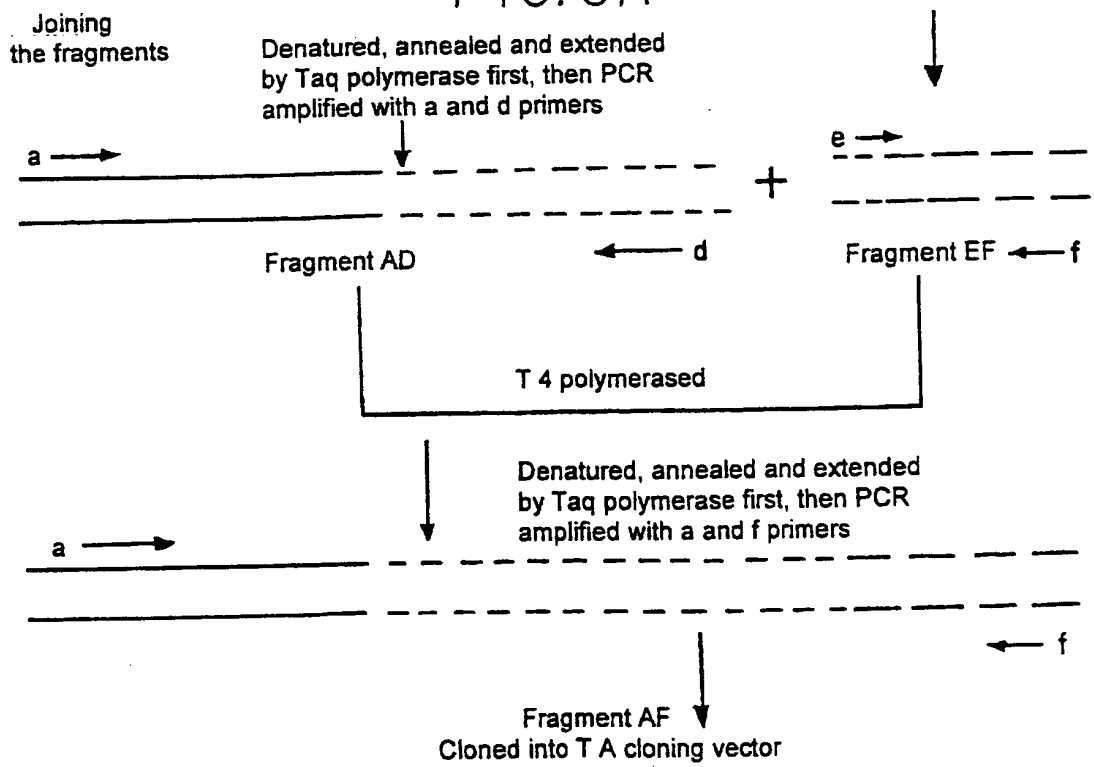


FIG. 6B

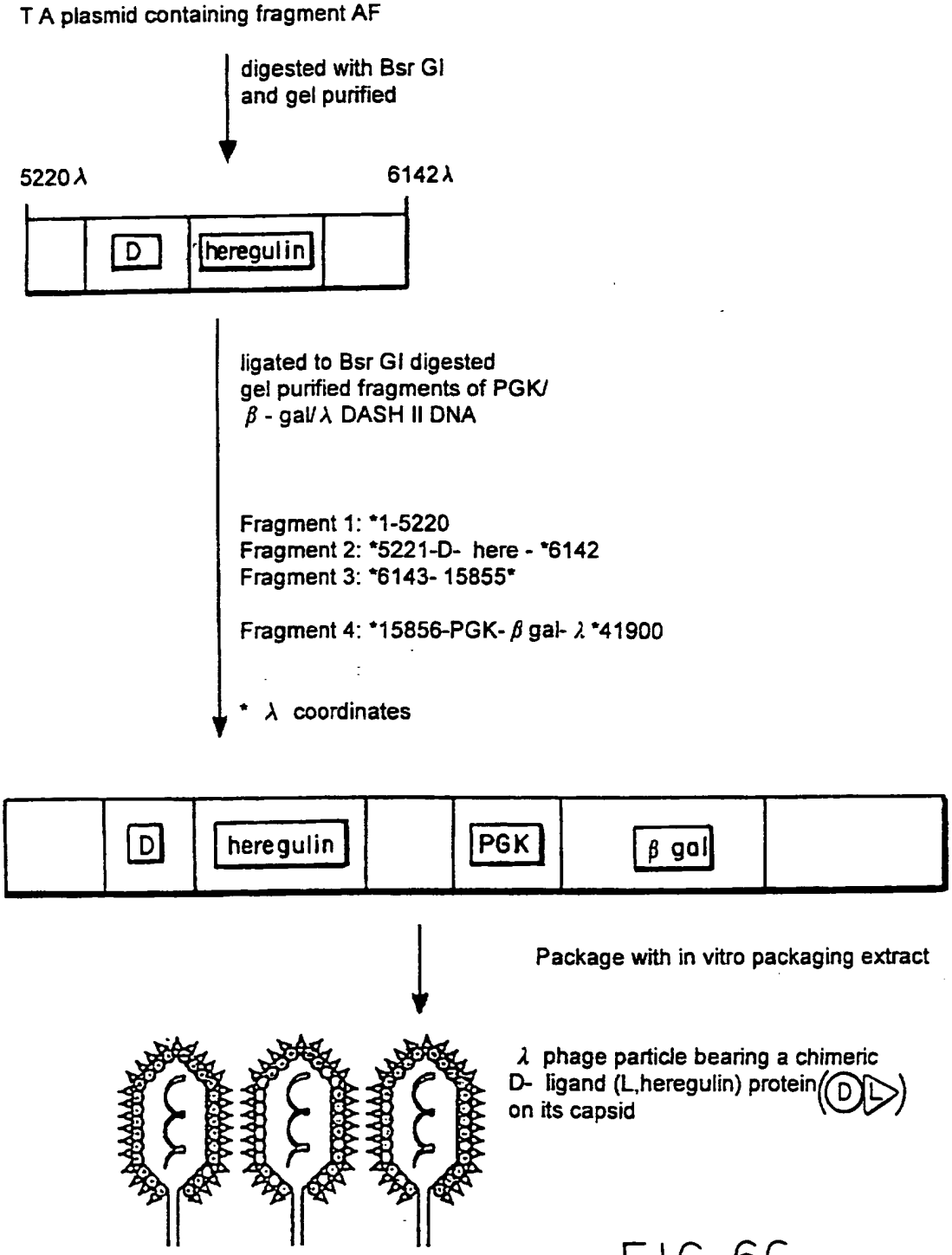

T A plasmid containing fragment AF
digested with Bsr GI
and gel purified
5220λ
6142λ
D
heregulin
ligated to Bsr GI digested
gel purified fragments of PGK/
β - gal/λ DASH II DNA
Fragment 1: *1-5220
Fragment 2: *5221-D- here - *6142
Fragment 3: *6143- 15855*
Fragment 4: *15856-PGK-β gal- λ *41900
* λ coordinates
D
heregulin
PGK
β gal
Package with in vitro packaging extract
λ phage particle bearing a chimeric
D- ligand (L,heregulin) protein
on its capsid
D
L

FIG. 6C

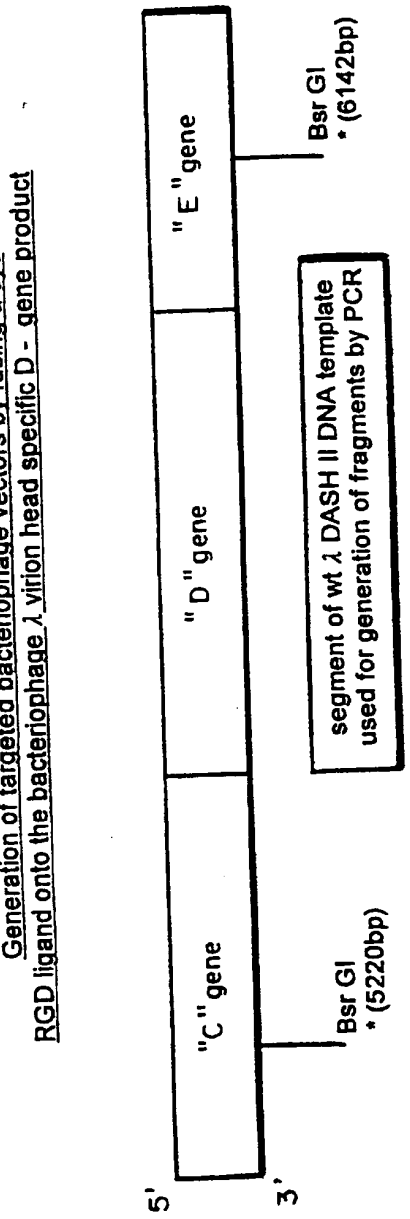

Generation of targeted bacteriophage vectors by fusing a cyc
RGD ligand onto the bacteriophage λ virion head specific D - gene product
"C" gene
"D" gene
"E" gene
Bsr GI
* (5220bp)
Bsr GI
* (6142bp)
segment of wt λ DASH II DNA template
used for generation of fragments by PCR
5'
3'

FIG. 7A

SUBSTITUTE SHEET (RULE 26)

Using specially designed primers (as shown below), the cyc RGD coding sequences were fused onto the 3' end of the λ DASH II "D" gene as follows:

Generation of fragments for gene fusion

description of primers
"a" contains only wt sequence of the λ DASH I "C": GENE.
"b" has a 3' end that is complementary to 3' end of the wt "D" gene and a 5' end which contains the coding sequence to cyclic RGD.
"c" has a 5' end that is complementary to the 5' end of cyclic RGD and a 3' end which is complementary to the 5' end of the wt "E" gene.
"d" contains only wt sequence of the λ DASH II "E" gene

PCR generation and amplification of fragments

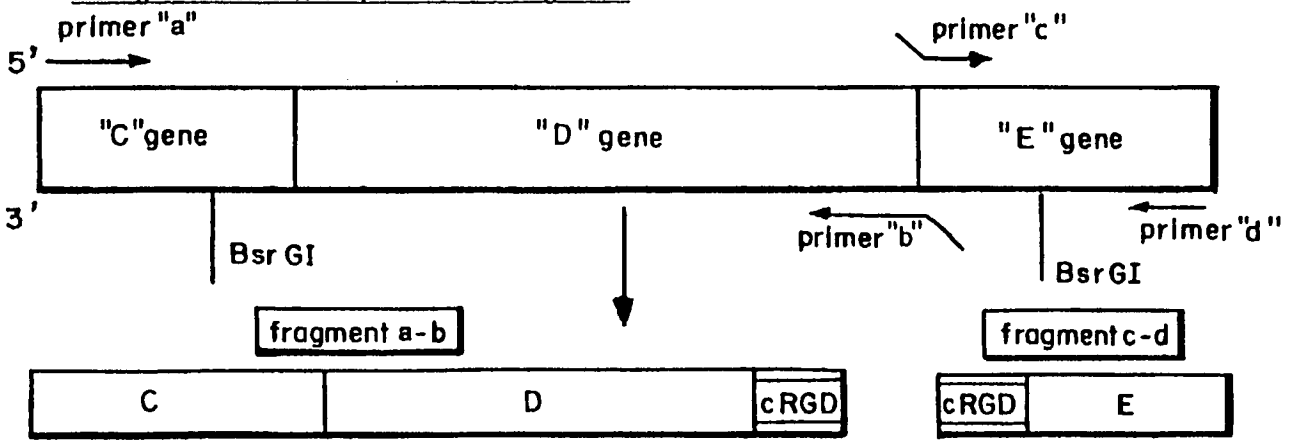


FIG. 7B

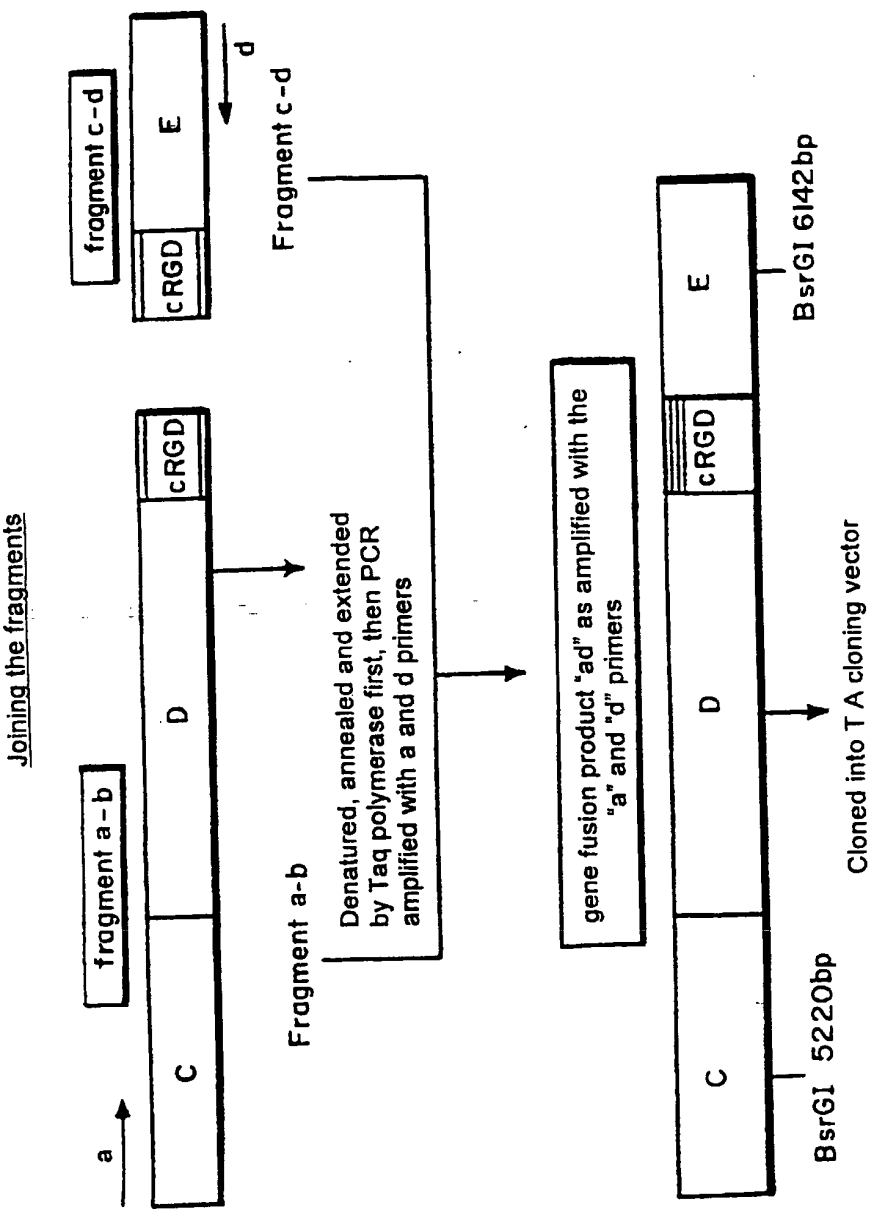

Joining the fragments
a
C
D
cRGD
fragment a-b
Fragment a-b
Denatured, annealed and extended by Taq polymerase first, then PCR amplified with a and d primers
fragment c-d
cRGD
E
d
Fragment c-d
gene fusion product "ad" as amplified with the "a" and "d" primers
C
D
cRGD
E
BsrGI 5220bp
BsrGI 6142bp
Cloned into T A cloning vector

FIG. 7C

Generation of targeted λ phage vector expressing the cyc RGD-D chimeric capsid and containing the CMV β Gal reporter gene

preparation of λ DASH II / CMV β Gal vector

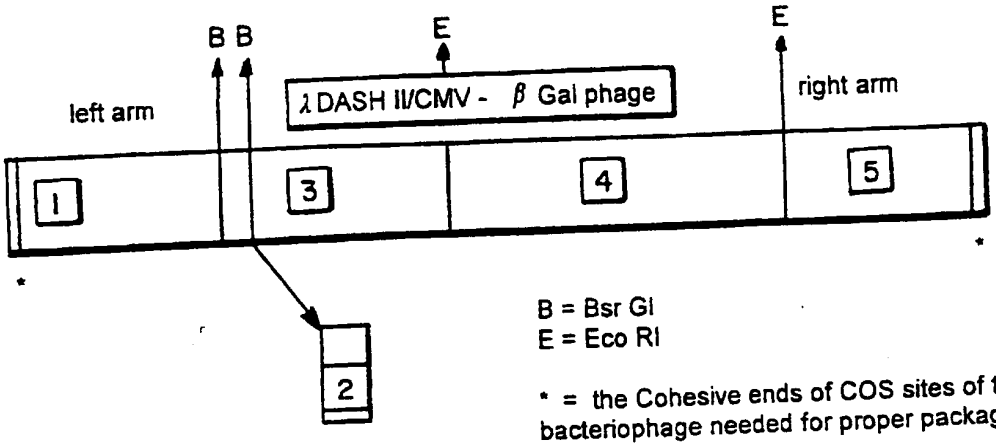
First, the λ DASH II / CMV β Gal bacteriophage vector was digested with Bsr GI and Eco RI* and cleaved into fragments. Fragment #2 (5220 - 6142bp) was then separated and removed by gel electrophoresis. (see below)

*co-digestion with Eco RI favors recombination of productive phage genome

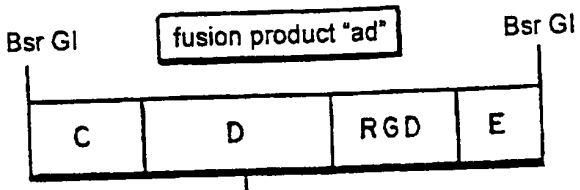
Fragment 1: *1 - 5220
Fragment 2: *5221 - *6142
Fragment 3: *6143 - 15855*
Fragment 4: *15856 - CMV - β gal - *41900
Fragment 5: *41900 - 48000
* λ coordinates

FIG. 7D-1

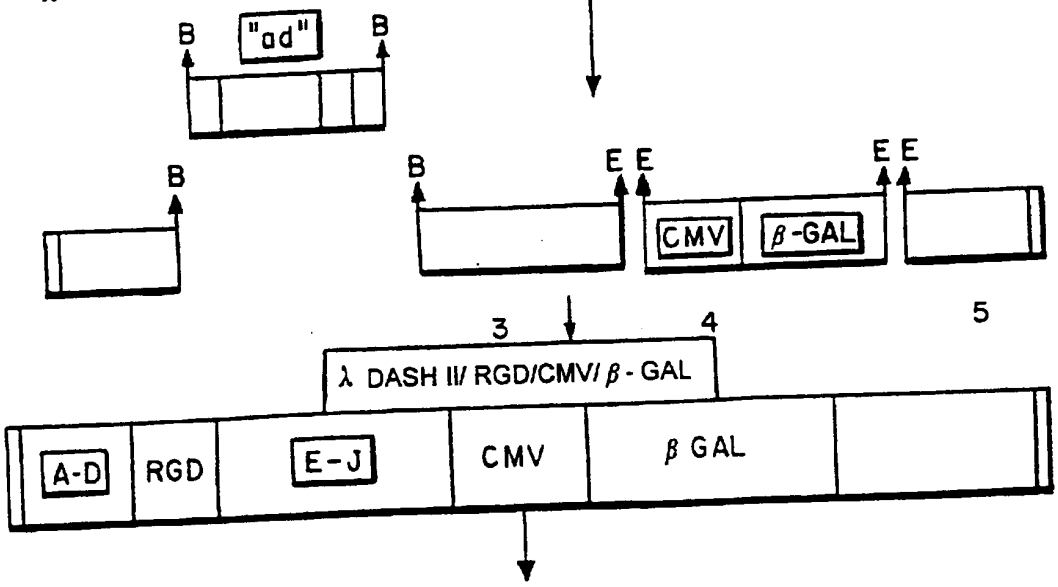
SUBSTITUTE SHEET (RULE 26)



Next, the TA plasmid containing the "ad" gene fusion fragment is digested with Bsr GI



λ DASH II/CMV β -Gal is ligated with the "ad" fusion product to form the
λ DASH II/RGD/CMV β -Gal phage construct



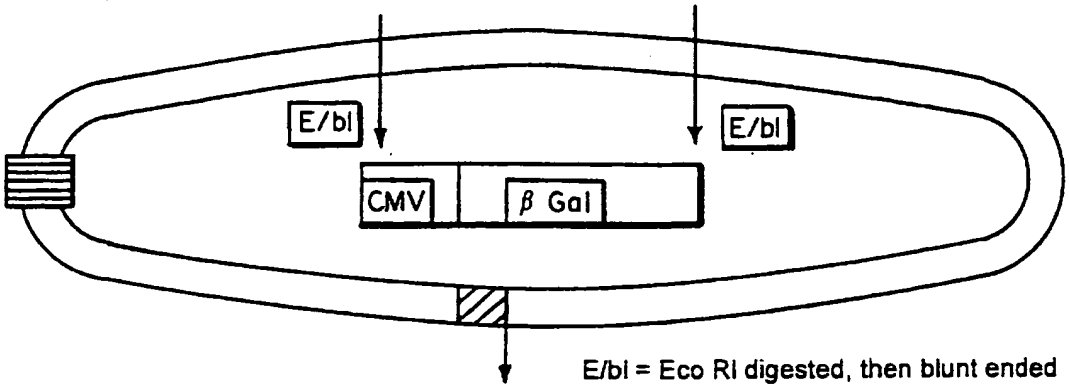
The ligated product was then packaged *in vitro* with *in vitro* packaging extract, the recombinant λ clones were characterized. Then λ DASHII/cyc RGD modified phage was then amplified and purified through CsCl banding

FIG. 7D-2

This parental λ DASH II/cyc RGD/CMV β Gal modified bacteriophage was subsequently used as a vector for targeted gene transfer of other DNA contructs as described below

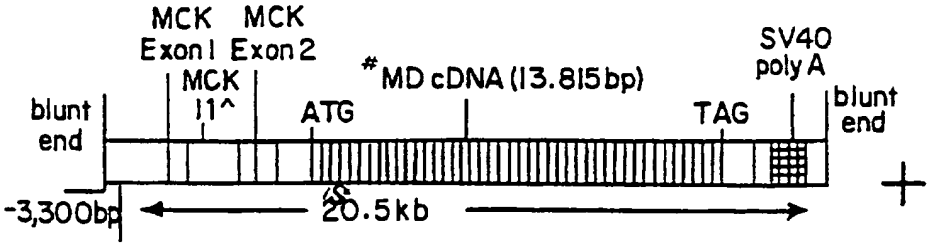
Generation of cyc RGD modified λ DASH II bacteriophage containing the murine dystrophin gene expression cassette

First, we ligated the CSO sites of the cyc RGD modified λ DASH II bacteriophage DNA together to protect them during the blunt ending process
Then the phage was digested with Eco RI and the CMV β Gal expression cassette was removed by gel electrophoresis



E/bl = Eco RI digested, then blunt ended

Murine dystophine gene expression cassette



pMDA PLASMID digested with BssHII and blunt ended

Ligated together with the blunt ended cyc RGD phage (ligated COS sites not shown)

* MCK P/E = muscle specific creatine kinase promoter/enhancer
# MD cDNA = Murine Dystrophine full length CDNA
^ MCK I1 = Muscle creation kinase intron 1

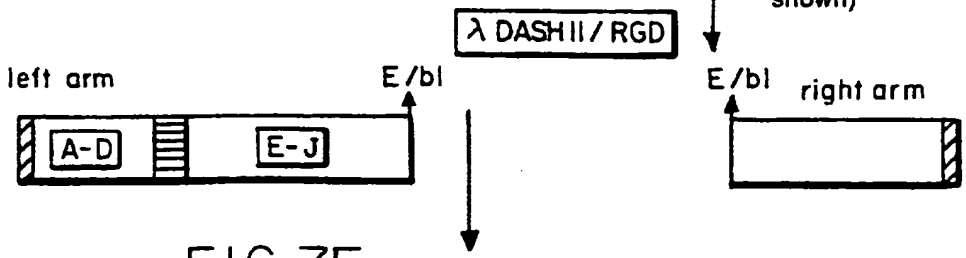


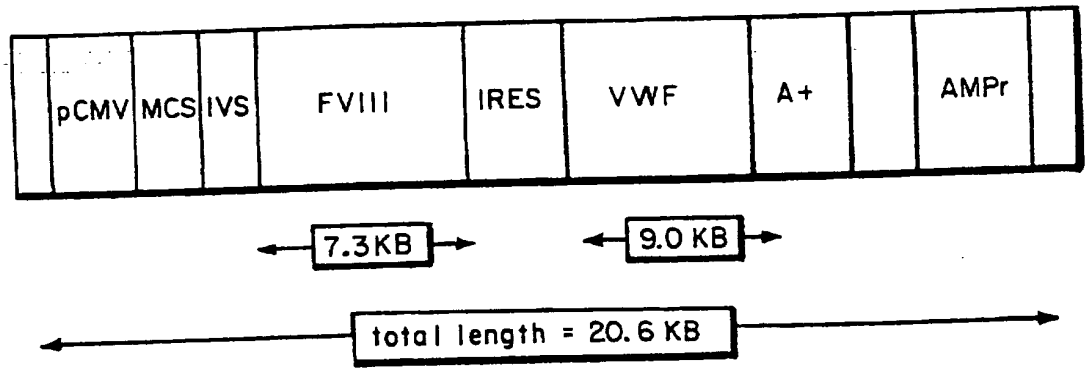
FIG. 7E

In vitro packaged, characterized, amplified and purified

Generation of cyc RGD modified λ DASH II bacteriophage containing the Factor VIII / IRES / Von Willebrand's Factor gene espression cassette

Construction of the plasmid containing the bicystronic espression cassette carrying the FVIII and the VWF genes

Gel purified, Eco RI digested, blunt ended, full length VWF cDNA was cloned into the Sma I site of the pIRES Δ *neo* (a derivative of the pIRES*neo* vector, in which the NEO gene was deleted by digestion with Kas I and Xba I and subsequent gel purification to remove the KasI / XbaI fragment. The vector was then blunt ended and religated to generate the pIRES Δ *neo* vector (the parental pIRES*neo* was obtained from Clontech, CA.) The resulting plasmid, pIRES Δ *neo* / VWF, was then digested with Eco RI, blunt ended, and ligated to Xho / Sal digested, blunt ended and gel purified full length factor VIII cDNA fragment. The resulting plasmid (as shown below) constitutes the bicystronic expression cassette containing the FVII I/ VWF genes.



Generation of the cyc RGD / λ DASH II / FVIII / IRES / VWF bacteriophage

Nru I digested pIRES Δ *neo* / FVIII / IRES / VWF plasmid DNA was used for ligation with the blunted left and right arms of the cyc RGD / λ DASH II prepared DNA. The ligated product was then *in vitro* packaged and clones were characterized, amplified and purified by CsCl banding.

FIG. 7F

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/12928

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :A61K 35/76; C12N 7/00, 7/01, 15/09, 15/85, 15/86
US CL :424/93.1, 93.6; 435/172.1, 172.3, 235.1, 320.1

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 424/93.1, 93.6; 435/172.1, 172.3, 235.1, 320.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, MEDLINE, EMBASE, BIOSIS, SCISEARCH, CAPLUS
search terms: bacteriophage#, phage#, eukaryotic, eucaryotic, mammal###, human, mouse, transfect###, transduc####, gene therapy

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 96/21007 (CHIRON VIAGENE, INC.) 11 July 1996 (11.07.96), page 4, line 30 to page 6, line 2, page 13, line 13 to page 22, line 14.	1-16
A	OKAYAMA et al. Bacteriophage lambda vector for transducing a cDNA clone library into mammalian cells. Molecular and Cellular Biology. May 1985, Vol. 5, No. 5, pages 1136-1142.	1-20
A	FLAMM et al. Expression of a cosmid containing the LCR, A-gamma, delta, and beta globin genes in mouse erythroleukemia cells. American Journal of Hematology. 1996, Vol. 53, pages 181-187.	1-20

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
02 SEPTEMBER 1997

Date of mailing of the international search report
24 SEP 1997

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231
Facsimile No. (703) 305-3230

Authorized officer
SCOTT D. PRIEBE
Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/12928

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	HART et al. Cell binding and internalization by filamentous phage displaying a cyclic Arg-Gly-Asp-containing peptide. Journal of Biological Chemistry. 29 April 1994, Vol. 269, No. 17, pages 12468-12474.	1-20